Transactions of the ASME®

Journal of Engineering for Gas Turbines and Power

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 122 • NUMBER 3 • JULY 2000

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published quarterly (Jan., April, July, Oct.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Engineering for Gas Turbines and Power, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or ... printed in its publications (B7.1, par. 3).

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048

(**Contents continued on inside back cover**)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ♾™
♻ 85% recycled content, including 10% post-consumer fibers.

Journal of Engineering for Gas Turbines and Power **JULY 2000**

Volume 122, Number 3

IN MEMORIAM

Professor James E. Peters
September 11, 1954–May 30, 2000

Professor James E. Peters was a devoted and dedicated faculty member in the Department of Mechanical and Industrial Engineering at the University of Illinois at Urbana-Champaign (UIUC) for close to twenty years. He received all his degrees from Purdue University — a Bachelor's degree in 1976, a Master's degree in 1978, and his Doctoral degree in 1981. Shortly after his graduation, Dr. Peters joined Illinois as an assistant professor and rose through the ranks to full professor in 1991. Professor Peters had been selected as the first James W. Bayne Professor and his appointment was being processed at the time of his passing. The University of Illinois at Urbana-Champaign awarded this professorship appointment to Professor Peters posthumously.

Professor Peters was a superb and dynamic teacher. He was exceptionally clear in his thinking and presentations, and he was extremely patient and helpful. He was a caring person and a dedicated teacher. Professor Peters earned national recognition for teaching as a recipient of the Society of Automotive Engineers Teetor Award. He was also honored with the highly coveted UIUC College of Engineering Everitt Award for Undergraduate Teaching Excellence, the highest award for teaching effectiveness in the College.

Professor Peters was a distinguished researcher in the field of combustion with primary emphasis on studies of two-phase flow phenomena. The overarching principles of his research were to examine complex problems from a fundamental viewpoint, and to use this fundamental knowledge to advance the understanding of important engineering applications. Through his more than ninety research publications, he provided fundamental contributions to the basic understanding of two-phase flows and spray phenomena including liquid spray behavior, the study of particulates for air pollution control, and coal combustion. In addition, his research has provided a long-lasting impact to significant applications in the areas of spark ignition engines, the use of alternative fuels, gas turbine performance, and particle diagnostics.

Professor Peters' research accomplishments were recognized within the university in 1986 when he was selected as a University Scholar. This honor recognizes and rewards outstanding scholars across all disciplines throughout the university. He gave numerous invited lectures, both nationally and internationally. He was a Fellow of the American Society of Mechanical Engineers (ASME).

In addition to his excellence in research and education, Professor Peters had a distinguished record of leadership within the Department, University, and Professional Societies. He served as the Associate Head for Graduate Programs, the Associate Head for Undergraduate Programs, and the Director of the Automotive Laboratory. He served on the UIUC Senate Committee on Admissions, the University Admissions Committee, Campus Committee on Financial Aid to Students, and Campus Senior Survey Committee. He served on and was the chair of the Campus Budget Oversight Committee, which provides budgetary leadership under the University's Budget Reform program. Professor Peters held leadership positions in various professional societies, including the Institute for Liquid Atomization and Spray Systems (ILASS), the Central States Section (CSS) of the Combustion Institute, and the ASME International Gas Turbine Institute (ASME-IGTI) Combustion and Fuels Committee. He served on the Board of Directors for both ILASS and CSS, and was the Chair of the CSS board. He was also the Chair of the Combustion and Fuels Committee. Professor Peters served as an Associate Editor for the ASME *Journal of Engineering for Gas Turbines and Power*.

Computational Simulation of Gas Turbines: Part 1—Foundations of Component-Based Models

John A. Reed
e-mail: jreed@memslab.eng.utoledo.edu

Abdollah A. Afjeh

Mechanical, Industrial, and Manufacturing Engineering Department,
The University of Toledo,
Toledo, OH 43606

Designing and developing new aerospace propulsion systems is time-consuming and expensive. Computational simulation is a promising means for alleviating this cost, but requires a flexible software simulation system capable of integrating advanced multidisciplinary and multifidelity analysis methods, dynamically constructing arbitrary simulation models, and distributing computationally complex tasks. To address these issues, we have developed Onyx, a Java-based object-oriented domain framework for aerospace propulsion system simulation. This paper presents the design of a common engineering model formalism for use in Onyx. This approach, which is based on hierarchical decomposition and standardized interfaces, provides a flexible component-based representation for gas turbine systems, subsystems and components. It allows new models to be composed programmatically or visually to form more complex models. Onyx's common engineering model also supports integration of a hierarchy of models which represent the system at differing levels of abstraction. Selection of a particular model is based on a number of criteria, including the level of detail needed, the objective of the simulation, the available knowledge, and given resources. The common engineering model approach is demonstrated by developing gas turbine component models which will be used to compose a gas turbine engine model in Part 2 of this paper. [S0742-4795(00)02303-6]

Keywords: Object-Oriented Modeling, Gas Turbines, Computational Simulation

1 Introduction

As the aerospace industry enters the 21st century, there is increasing pressure to reduce the time, cost, and risk of aircraft engine development. To compete effectively in today's global marketplace, innovative approaches to reducing propulsion system design-cycle times are needed. An opportunity exists to reduce design and development costs by replacing some of the large-scale testing currently required for product development with *computational simulation* [1]. Increased use of computational simulation promises not only to reduce the need for testing, but also to enable the rapid and relatively inexpensive evaluation of alternative designs earlier in the design process [2].

The traditional method for numerical simulation of the gas turbine engine decomposes the system into an assembly of component models. Component operation is defined using appropriate conservation equations, and the entire engine is simulated by solving the resulting system of equations. In the past, limited computer resources forced component representations to be based on relatively simple equations and empirical data obtained from component and system testing [3–6]. Over the years, more detailed numerical models based on first principles have emerged. These models utilize advanced discrete numerical methods such as Computational Fluid Dynamics (CFD) and Finite Element Analysis (FEA), to allow more realistic and detailed investigations of fluid and mechanical process occurring in engine components. Integration of high-level CFD and FEA-based models into system-level engine models can improve design cycle times by reducing the need to obtain model data experimentally. However, the process of integrating advanced high-level numerical methods into existing system-level engine simulation software is difficult. Current system-level gas turbine simulation programs, incorporating single-disciplinary models and designed to operate on single processor machines, are not sufficiently robust and flexible to support integration of detailed CFD and FEA analyses [7]. As a result, gas turbine engine simulation software still rely heavily on empirical models.

The purpose of this two-part paper is to describe a new object-oriented computational simulation framework we have developed in cooperative research with NASA Lewis Research Center's NPSS (Numerical Propulsion System Simulation) project [8]. The framework, called *Onyx*, is capable of integrating advanced numerical methods for detailed interdisciplinary analysis of complete gas turbine engines at various levels of fidelity. In Part 1, we present Onyx's Common Engineering Model (CEM), which forms the foundation for an integrated representation of the engine, it's components, subcomponents, and subassemblies. A frequent deficiency in modeling and simulation of multidisciplinary systems is the lack of a suitable representation of the analytical and geometric models. Onyx's CEM can accommodate models having differing fidelity and discipline, and it's internal data structure is capable of providing direct access to CAD geometry for mesh generation and visualization software tools. The CEM also provides a general mechanism to couple data between different disciplines and fidelities. Furthermore, the CEM has provisions for integrating legacy software packages within the model for inclusion in the overall simulation. The CEM design is implemented using the Java™ object-oriented programming language [9].

Part 2 describes the Onyx simulation framework. Advanced gas turbine simulation requires a distributed simulation environment which is flexible, customizable, extensible and user-friendly. Onyx is designed to allow the user to customize simulation functionality by allowing components to be replaced by other components having different functionality. Such ''pluggability'' is essential for keeping the architecture current. Similarly, it is capable of supporting the integration of new simulation techniques and computing methodologies with as little effort as possible. The introduction of interdisciplinary models and models having differing levels of fidelity requires support for distributed computing as

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-346. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

it cannot be assumed that the higher-fidelity software will run efficiently (or at all) on the same computer platform as the rest of the system. Onyx supports distributed computing using several widely-used software distribution mechanisms, such as CORBA (Common Object Request Broker Architecture) and RMI (Remote Method Invocation). Finally, Onyx provides a highly portable graphical user interface to reduce the efforts of developing new models and executing simulations.

This paper is organized as follows: the following section briefly introduces the methodology of object-oriented modeling and simulation. Section 2 presents the idea of hierarchical decomposition, and its application to gas turbine engines. Section 3 describes the conceptual foundations for the common engineering model, while Section 4 provides the object-oriented class implementation of the model. As a representative example of the Common Engineering Model development process, a transient, lumped-parameter, aero-thermodynamic compressor component model, is presented in Section 5.

2 Object-Oriented Modeling, Simulation

Computer-based simulation involves designing a model of an actual or theoretical physical system, executing the model on digital computer, and analyzing the execution output [10]. Models are generally developed by defining a given problem domain, reducing the physical entities and phenomena in that domain to idealized form based on a desired level of abstraction, and formulating a mathematical model through the application of conservative laws. To execute the model on computer, a suitable software program is designed which implements the appropriate computational methods to solve the mathematical model.

Traditionally, simulation software development has followed the top-down structured design approach, which applies the method of *functional decomposition* to establish program structure. In this method, the program is decomposed on the basis of functionality, with each module (e.g, FORTRAN subroutines) corresponding to a step in the overall algorithm [11]. Although this method is theoretically sound, and has produced the majority of software systems to date, it fails to facilitate the implementation of software models which reflect the structure of the real-world, and are flexible and reusable in all aspects of gas turbine simulation. Object-oriented design methods, in which emphasis is placed on modeling systems as collections of interacting *objects*, rather than *processes*, provide an improved approach to constructing models for gas turbine simulation.

An object is a software package containing data (variables) and the operations (methods) which operate on that data. An object performs an operation when it receives a request (message) from another object. Methods are described through well-defined *interfaces* which provide information concerning the object's variables and methods that are visible to clients of the object. The interface does not define how an object implements the operation, but only to what messages the object responds. Sending a message to an object through its interface to execute an operation, is the only way to get an object to change its internal data. Consequently, an object's implementation is *encapsulated*—hidden from client objects.

An object's implementation is defined by its *class*. A class is a template that specifies the object's internal data representation, its methods, and their implementation. Objects are created by *instantiating* a class; the object is said to be an *instance* of the class. The class defines characteristics shared by all instances of the class, whereas each instance (i.e., object) of the class contains internal data, known as instance variables, which make it unique. New classes are defined in terms of existing classes using *class inheritance*. When one class is a *subclass* of another, it is said to inherit the representation and behavior of its *parent* class. Objects which are instances of the subclass contain all variables defined by both the subclass and the parent class; and they inherit the ability to perform the methods defined by both the subclass and parent class. Subclasses may *override* a method defined by their parent class. Overriding allows a subclass to handle a message instead of the parent class. In this manner, subclasses may extend and redefine the behavior of the parent class. Such specialization provides a natural mechanism for representing hierarchical systems.

Class inheritance is a fundamental feature found in most object-oriented programming languages. Well-designed class hierarchies allow for reuse of abstraction, design, and functionality, defined in parent classes. Inheritance, it is argued, lets programmers improve productivity by developing new classes that differ slightly from existing classes by specializing the existing classes, rather than starting from scratch. However, developing large software systems by reuse through inheritance is often difficult. One reason is that parent classes typically define at least part of their subclass' physical representation. Since inheritance exposes the internal representation of a parent class to its subclass, it is said to break encapsulation [12]. Because a subclass' definition is tied to the representation of its parent class, any change in the parent class representation will affect the subclass.

Object composition is an alternative to class inheritance. With composition, new functionality is obtained by assembling an object from other objects. An important design requirement of object composition is that the objects being composed have well-defined interfaces which allows one object to work with other objects. Because objects are accessed only through their interfaces, encapsulation is not broken. Most importantly, an object may be replaced (at run-time) by another object as long as they share the same type.

3 Hierarchical Model Abstraction

Structure. Object composition provides a powerful mechanism for representing the physical topology of the gas turbine engine. From a structural viewpoint, a gas turbine engine is essentially an assembly of *engine components*—inlet, fan, compressors, combustor, duct, turbine, shafts, and nozzle (see Fig. 1(*a*)). These components can be represented in the computational domain as objects. If the interfaces between the component objects are correctly specified, they may be combined without restriction to form various derivative propulsion systems, such as multi-spool or turbo-prop engines.

Composition may also be employed recursively to represent the hierarchical structure of a model. Consider the Fan component shown in Fig. 1(*a*). The physical structure of the Fan may be decomposed into its constituent parts; these may include: hub, casing splitter, flow-field and stage objects (see Fig. 1(*b*)). Each of these sub-components can be represented in the software model as an object, and aggregated to form a Fan composite object. The process, can be extended as desired based on the level of resolution desired in the given model. The stage object, for example, may be further decomposed into rotor blade and stator blade objects.

Component models are thus organized in a hierarchical structure which reflects the actual physical structure of engine, its subsystems and components. Component models at the highest level of the hierarchy show the least detail; and each succeeding level reveals more detail until a fundamental set of sub-component objects is reached. These primitive component object's behavior is described in terms of mathematical equations.

Behavior. Simulating complex systems requires the development of a hierarchy of models, or *multimodel*, which represent the system at differing levels of abstraction [13]. Selection of a particular model is based on a number of (possibly conflicting) criteria, including the level of detail needed, the objective of the simulation, the available knowledge, and given resources. For preliminary gas turbine engine design, simulation models are often used to determine an engine's thrust, fuel consumption rates, etc. These simulations generally use relatively simple zero-dimensional component models to predict performance. However, in other situations, such as *multidisciplinary analysis*, higher-

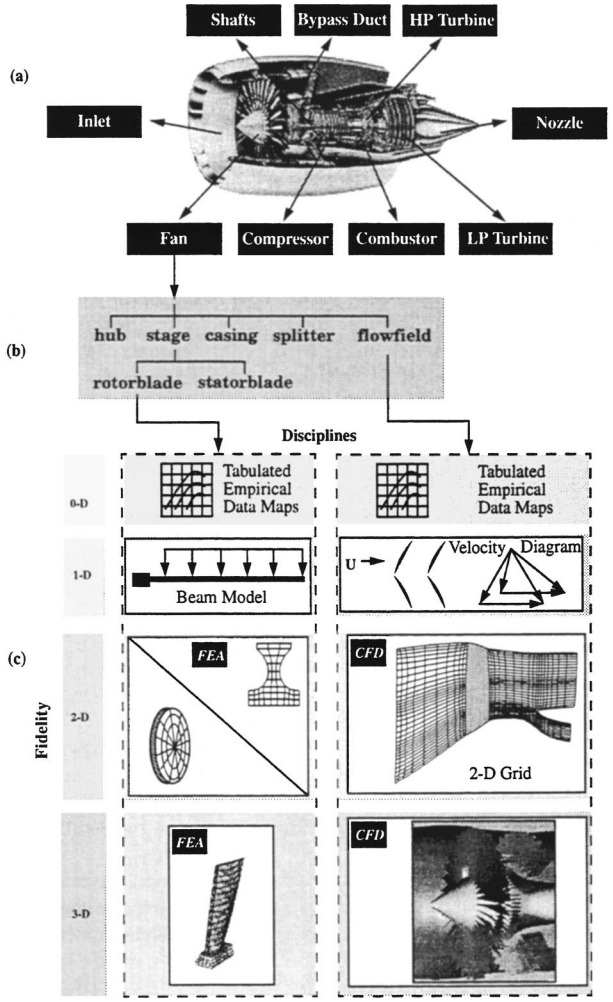


Fig. 1 Structural and behavioral abstraction of gas turbine engine and it's components

order models are needed. For example, to prevent the possibility of a fan blade rubbing the cowling, an engineer might perform a coupled aerodynamic, thermal and structural analysis of the blade to determine the amount of blade bending due to the thermal and aerodynamic loading. Such an analysis would require several high-fidelity models using fully three-dimensional, Navier-Stokes CFD, and structural FEA algorithms.

The conceptual view of a gas turbine multimodel is shown in Fig. 1(*c*). Here, multiple views of the Fan's rotor blade and flow-field objects are depicted. Moving down the figure, the spatial resolution (and thus complexity) of the views increases from zero-dimensional to three-dimensional. The rotor blade is represented using various mechanical/structural methods, while the flow-field is represented by various aero- and fluid-dynamic methods.

Zero-Dimensional. At the zero-dimensional level, component objects are modeled using space-averaged algebraic and/or ordinary differential equations, performance maps, and design geometry to provide operating characteristics for both steady-state and transient operations. Fluid dynamic processes are generally represented using isentropic relations and baseline map objects which define relationships between stagnation pressure, rotor speed, efficiency, and mass flow rate [14]. Additional maps are also utilized to account for secondary effects due to stator blade settings as well as structural characteristics such as bending.

One-Dimensional. A one-dimensional view of the Fan fluid flow is typically based on a row/stage model using velocity diagrams [15]. Structural effects might be defined using a simple beam model to determine blade displacement and frequency responses.

Two-Dimensional. The two-dimensional view of the flow is represented by an axi-symmetric CFD model. Blade force, loss and blockage terms, arising from circumferential averaging of boundary conditions in the full Navier-Stokes equations, are evaluated from either experimental component testing or three-dimensional component CFD simulations [16]. Structural analysis may be performed on axial or disk finite-element mesh models using structural analysis software, such as NASTRAN.

Three-Dimensional. The three-dimensional view of the flow is generally represented by a Reynolds-averaged Navier-Stokes solution on a finite-volume mesh representing a single blade passage. The structural analysis is performed on a single rotor blade model using a finite-element mesh to discretize the mechanical equations.

4 Common Engineering Model

In Onyx's Common Engineering Model, physical engine component structures and substructures are represented by *EngElement*, *Domain Model*, *Port*, and *Connector* objects. The process of identifying, abstracting and mapping from the physical domain to these objects in the computational domain is illustrated in Fig. 2. Consider the Inlet and Fan components in a turbofan aircraft engine (Fig. 2(*a*)). Based on a functional decomposition via control volume analysis, separate control volumes are identified for the Inlet and Fan (Fig. 2(*b*)). Using symmetry, the control volumes are defined from the engine centerline into the outer cowling. The two control volume boundaries—identified by the dashed lines—share a common boundary or control surface located between the Inlet and Fan. At this interface, matter and energy are exchanged between the two components. Component behavior is specified by conservation laws which yield sets of

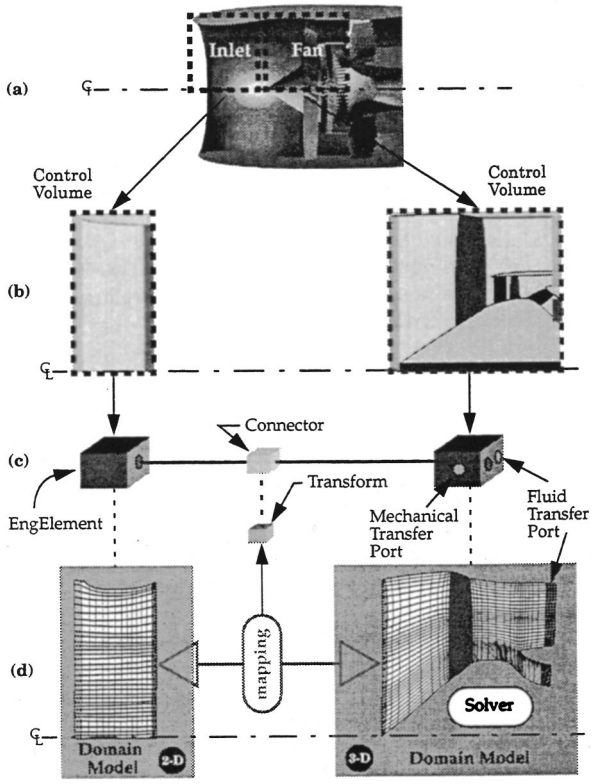


Fig. 2 Conceptual mapping of physical gas turbine domain to Onyx computational framework objects

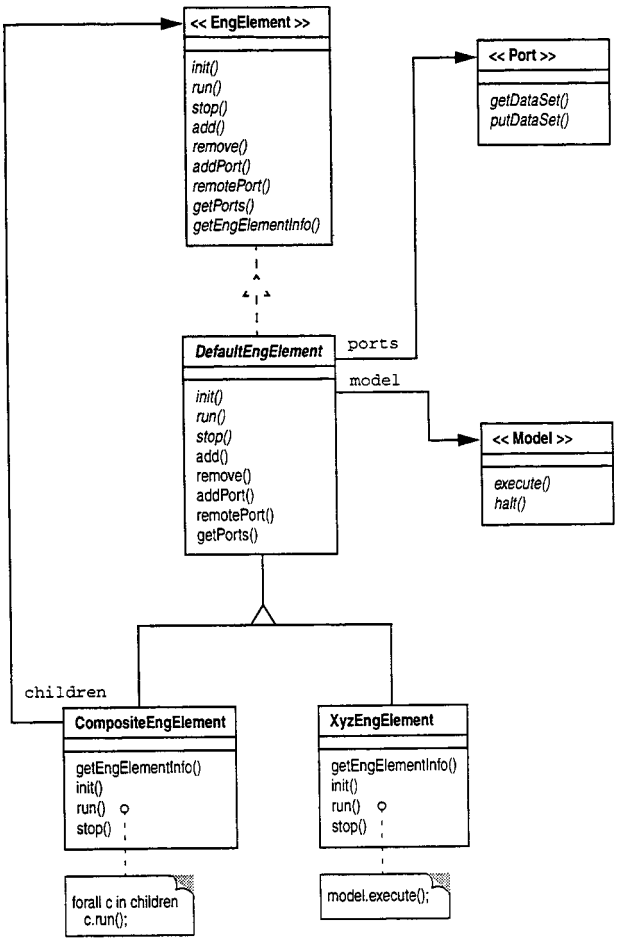


Fig. 3 EngElement class structure diagram

equations describing the physical processes occurring in the volume. For an aero-thermodynamic analysis, conservation relationships for continuity, momentum and energy are typically used [17].

EngElement, Domain Model, Port and Connector descriptions are represented in Onyx by Java *interfaces*. A Java interface contains no code; rather, it defines a set of method signatures that *must* be defined by any class that *implements* the interface. Classes which implement the interface define the specific behavior for each method, and can then be subclassed to define more specialized classes. Java interfaces provide a powerful mechanism for describing design intent without restricting it to a particular implementation. A variable whose type is an interface may refer to any object whose class implements the interface. Interfaces are identified by the ''«interfaceName»'' notation in the class diagrams (for example, see the EngElement interface in Fig. 3).

4.1 EngElement. The abstract nature of the control volume provides a convenient and generic representation for objects in the computational domain. In Onyx, *EngElement* objects are used to represent the physical control volume (see Fig. 2(*c*)). It should be noted that, in this example, the control volumes represent high-level components; but there is no requirement that this always be the case. The EngElement class structure (see Fig. 3) is based on the *Composite* design pattern [18]. This pattern effectively captures the part-whole hierarchical structure of the gas turbine component models. The advantage of the Composite pattern is that it allows the uniform treatment of both individual objects and compositions of objects. This is essential in providing a generalized, modular and flexible modeling system.

EngElement is a Java interface which establishes the key behavior for all engine component classes incorporated into Onyx. It declares the core methods needed to initialize, run and stop engine model execution, as well as methods for managing attached Port objects. The abstract class DefaultEngElement implements EngElement and provides default functionality for the interface methods. It maintains a list of Port objects associated with the EngElement as well as a reference to its Domain Model object. References to these objects are indicated in the figure by the variable names, `ports` and `model`. Note that these variables reference Java interfaces, thus allowing EngElement to reference any object which implements either of those interfaces. This is essential for providing plug-compatible models and provides the foundation for model composition in Onyx.

In most cases, developers will subclass DefaultEngElement to create concrete engine component classes, such as class XyzEngElement, to implement the required functional methods. Subclasses of DefaultEngElement are required to implement concrete methods for `init`, `run`, and `stop`. These methods generally delegate their operation to the `model` object. The advantage of delegating execution to the Domain Model is described in the next section.

The design approach of providing a default abstract class for a Java interface is used throughout the Onyx framework to give the developer more flexibility when plugging in new classes. Here, the developer may select to inherit the functionality provided by DefaultEngElement, or to inherit from another class and implement the methods defined by the EngElement interface. In both cases, the new class would be of type EngElement, and thus could be plugged into the Onyx framework.

CompositeEngElement defines an aggregate of EngElements, which is stored in a list and referenced by the `children` variable. CompositeEngElement implements `init`, `run`, and `stop` which calls those methods on each of its EngElement children. Management operations for children are declared in DefaultEngElement to maximize component transparency. To ensure type-safety, these methods throw an exception for illegal operations, such as attempting to add or remove an EngElement from another EngElement, rather than a CompositeEngElement.

4.2 Domain Model. As indicated above, primitive component behavior is defined using mathematical equations. Commonly, these equations and accompanying boundary conditions are so complex that analytical solutions cannot be found. Approximate solutions are obtained by discretizing the equations on a geometrical mesh, and solving the equations numerically. Figure 2(*d*) illustrates grid geometry for both the Inlet and Fan components. The mathematical equations, grid geometry and solver are encapsulated by a *Domain Model* object in Onyx.

The complexity of the various Domain Models suggest that it is desirable to encapsulate the Domain Model and remove it from the structure of EngElement (see Fig. 2(*c*)). This enhances the modularity of EngElement, allowing new EngElement classes to be added without regard to the Domain Model, and conversely to add new Domain Models without affecting the EngElement class. To achieve this, the CEM utilizes the *Strategy* design pattern [18] to encapsulate the Domain Model in a separate object. The benefit of this pattern is that families of similar algorithms become interchangeable, allowing the algorithm—in this case the Domain Model—to vary independently from the EngElements that use it.

Furthermore, by encapsulating the Domain Model in a separate object, this design encourages the ''wrapping'' of pre-existing, external software packages. For example, the Fan Domain Model in Fig. 2(*d*) might ''wrap'' the ADPAC three-dimensional flow solver [19] to provide steady-state aerodynamic analysis of fluid flow within the Fan. This approach allows proven functionality of existing software analysis packages to be easily integrated within an EngElement.

The Domain Model class structure (see Fig. 4) is designed to be very general, due to the complicated nature of the various models

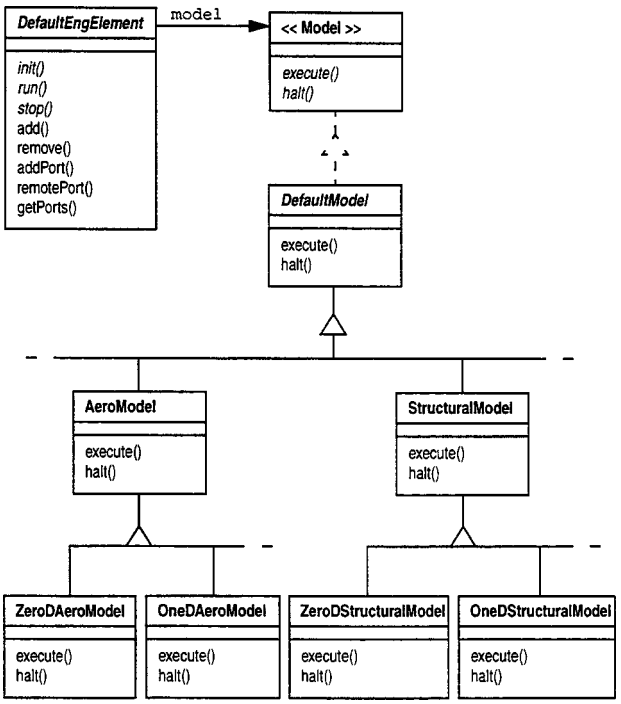


Fig. 4 Domain Model class structure diagram

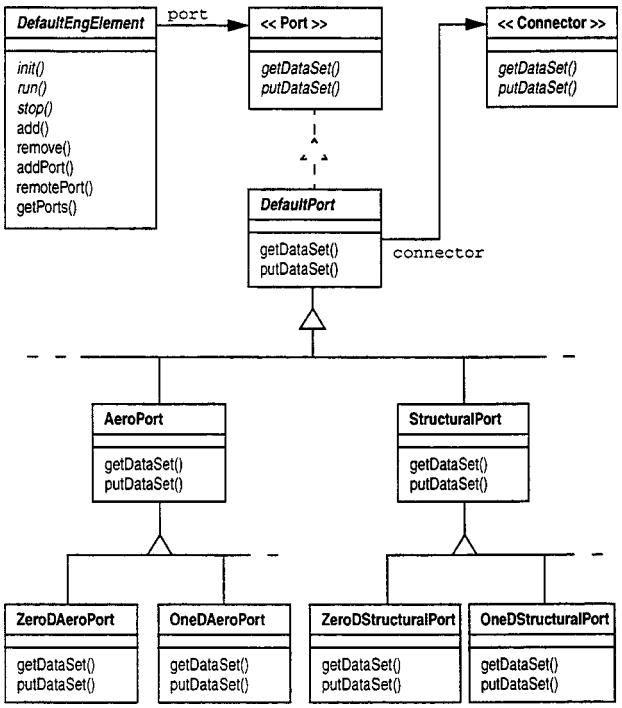


Fig. 5 Port class structure diagram

which can be encapsulated in an EngElement. The intent is not to restrict the use of any algorithm or the ''wrapping'' of external software packages by overly-defining the Model interface. Consequently, Model defines only two methods, `execute` and `halt`, which is used to start and stop the execution of the Domain Model code. Additional methods are obviously needed to access and make the data internal to the Domain Model available to the EnggElement, but because these are specific to the Domain Model structure, they are not included in the Model interface.

The DefaultModel class is an abstract class which implements the Model interface. For illustrative purposes, a hierarchy of Model subclasses is shown. Subclasses are first arranged hierarchically by Domain Model discipline (e.g., aerodynamic, structural, etc.). The discipline-specific subclasses may be further subclassed by fidelity level (e.g., zero-dimensional, one-dimensional, etc.). This diagram only serves to illustrate one possible hierarchy; many other alternative structures could be developed. This reinforces the design intent of a general class structure with which to encapsulate various Domain Models.

4.3 Port. EngElements have zero or more *Port* objects associated with them. Ports represent a surface on a control volume through which some entity passes (e.g., mass or energy). Ports are generally classified by the entity being transported across the control surface. In this example, the Inlet has two fluid transfer Ports—one at entrance to the Inlet and the other at the exit. These are identified by lines in the control volume figure (Fig. 2(*b*)), and by circles on the EngElement surface in Fig. 2(*c*). The Fan has three fluid transfer Ports: one at the entrance to the Fan which receives fluid from the Inlet; one at the exit for fluid entering the Bypass Duct; and the third for fluid entering the Compressor. The Fan also has a mechanical transfer port, which defines the control surface on the Fan through which mechanical energy is passed (in this case from the Low-speed Shaft).

The Port class structure is shown in Fig. 5. The Port interface defines two methods to set and retrieve the data defined by the Port. Class DefaultPort defines default functionality for these methods, and maintains a reference to the Connector object currently connected to the Port. As with the Model class hierarchy, the Port class structure is extremely general with subclasses of the DefaultPort class being defined as required. The subclasses shown in Fig. 5 are subclassed according to the discipline (e.g., aerodynamic, structural, etc.), and those classes are then each subclassed by fidelity (zero-dimensional, one-dimensional, etc.).

The specific subclass of Port an EngElement instantiates is determined by the discipline-fidelity combination of the EngElements Domain Model(s). For example, if EngElement has a Domain Model which is a zero-dimensional, aerodynamic model, then an instance of EngElement creates ZeroDAeroPort objects to handle input and output. Because EngElement delegates its behavior to the Domain Model object, the selection of the model is dynamic and may be changed at run-time. Consequently, the Port objects, which are inherently related to the Domain Model must also change accordingly. The dynamic creation and management of Ports is represented by the *State* design pattern [18]. This pattern allows the `port` object, in DefaultEngElement, to appear to change type as the fidelity and discipline of the Domain Model changes.

The strength of the Port class structure comes from its generality. As with the Model class structure, the intent was not to overdesign the class and impose rigid constraints on developers. While there is a strong coupling between Ports and Domain Model objects due to the underlying format used to store the data, the Port class minimizes this coupling by defining a simple, consistent interface with which to access the control surface data. This interface allows Connector objects to connect to any object as long as it is of type Port, making Port objects pluggable to Connectors.

4.4 Connector. The interface between successive connected control volumes is represented in the computational domain by *Connector* objects. Because the interface represents an exchange of information (e.g., mass, energy, etc.) between two control surfaces, its representation in Onyx is defined by a connection between two Ports. This is shown in Fig. 2(*c*).

A consequence of allowing multifidelity and multidisciplinary models to be incorporated in an engine component is that an engine model may be composed of component models having differing fidelity and/or discipline. While this is attractive in order to perform *component zooming* [20] to reduce computational re-

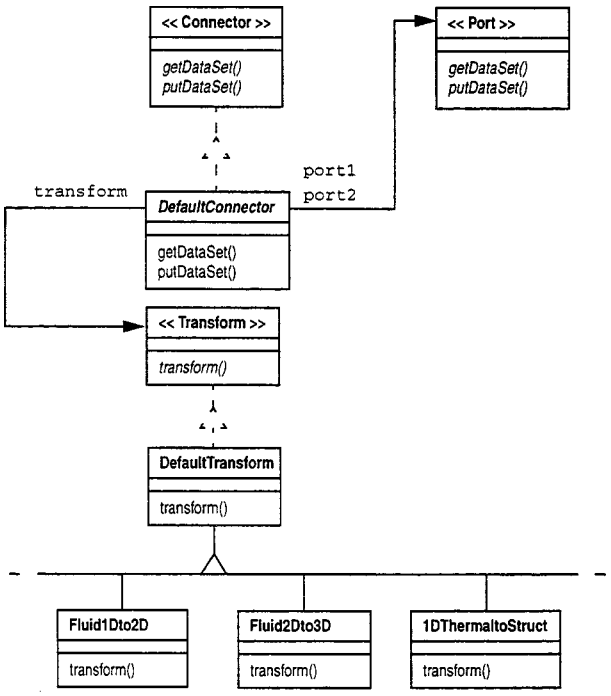


Fig. 6 Connector class structure diagram

quirements of the simulation, it introduces difficulties in passing data from one component to another. Referring to Fig. 2(*d*), if the Inlet component's behavior is defined by a two-dimensional fluid model, while the Fan is represented by a three-dimensional fluid model, there is a mismatch in fidelity. To pass data from the Fan to the Inlet requires some intelligence capable of transforming three-dimensional Fan data into a two-dimensional Inlet data; and vice-versa for data exchange from Inlet to Fan. A similar case occurs in multidisciplinary analysis where disciplinary coupling is performed and the results from one disciplinary analysis must be mapped onto the computational domain of another discipline. For example, mapping structural analysis results from a finite element mesh to a finite volume mesh used for aerodynamic analysis. In these cases, some intelligent process is needed to correctly map data from one discipline domain to another.

Holt and Phillips [21] introduced the concept of *connector* objects to provide appropriate methods for ''expanding'' or ''contracting'' the data, and mapping from different discipline domains. In Onyx, the concept is refined to improve reusability. Specifically, the data transformation is delegated to a separate *transform* object (see Fig. 2(*c*)). The Connector class structure is shown in Fig. 6. The Java interface, Connector, defines the core interface functionality. As with previous interfaces, an abstract class, DefaultConnector, which implements the interface, provides default implementation of each method, and defines the variables `port1` and `port2`. These variables reference the Port objects which are connected to the Connector object. In order to enhance reusability of the various transformation strategies, that responsibility is delegated to a separate Transform object. Transform objects then, encapsulate the necessary intelligence to expand/contract data and map data across disciplines. The Connector class structure employs the *Strategy* pattern [18] to define a family of Transform classes, each encapsulating an appropriate algorithm to expand, contract or map data. Figure 6 depicts this family of classes as subclasses of the abstract DefaultTransform class. DefaultTransform implements the Transform Java interface which defines a single method, *transform*, used to perform the data transformation. The *Strategy* pattern allows the different transformation algorithm classes, such as Fluid1Dto2D, Fluid2Dto3D, and 1DThermaltoStruct to be interchangeable within a Connector. Connector selects an appropriate Transform object using the *State* pattern [18], based on the fidelity-discipline combination of the connection. This pattern allows the *transform* object to alter its behavior as needed to accommodate the given fidelity-discipline combination.

5 Example

In this section, we illustrate the process of defining component models for gas turbine simulation in the Onyx framework. The models are developed following the methodologies outlined in the previous sections. In Part 2 of this paper, these models will be graphically assembled in Onyx's Visual Assembly Framework, and the engine model simulated.

The example engine component models are based on component designs for the NASA/General Electric Energy Efficient Engine (E^3) [22]. Designed and built in the early 1980's to research the advancement of high-bypass commercial turbofan engines, the E^3 is representative of modern turbofan engines, such as the GE90 engine (shown in Fig. 1(*a*)). The E^3 was selected for study due to the availability of non-proprietary component test rig data, and assembled engine performance data.

Following the control-volume approach outlined in Section 4, the engine was decomposed into its main components: Inlet, Fan, Bypass Duct, Compressor Combustor, High Speed Shaft, Low Speed Shaft, High Pressure Turbine, Low Pressure Turbine Nozzle, and a series of Bleed Ducts, which carry cooling air from the Compressor to the Combustor, High-Pressure Turbine, Low-Pressure Turbine, and provide customer air for driving equipment on the aircraft. The following set of EngElement classes were created to represent each of these physical components: Inlet, Bleed, Combustor, Compressor, Duct, Environment, Nozzle, Shaft, and Turbine. In addition, a MixingVolume class was added to provide dynamic modeling capability. The MixingVolume represents an inter-component volume, which is placed between successive, non-MixingVolume EngElements.

Following the mathematical development given in Daniele et al. [23], zero-dimensional, aero-thermodynamic Domain Models were written for each EngElement. Rather than describe the implementation of each EngElement class (which would be very lengthy), an illustrative example is presented based on the development of the Compressor class. Each of the other EngElement classes used in this example were developed in a similar fashion and have a similar structure; a complete description can be found in Reed [24].

5.1 Compressor. The Compressor EngElement represents a variable-stator compressor used to increase the kinetic energy of the working fluid in the compressor by transmitting the mechanical energy from a shaft to the fluid. The Compressor EngElement is also used to model the Fan component in the example engine model.

Compressor performance is represented by a set of overall performance maps normalized to design point values. Baseline performance maps provide normalized inlet-corrected mass flow rate and normalized efficiency as a function of normalized pressure ratio and normalized inlet-corrected spool speed. Shifts in normalized inlet-corrected mass flow rate, based on (un-normalized) off-schedule values of variable stator position, are also provided as a function of normalized pressure ratio and normalized inlet-corrected spool speed.

Normalized inlet-corrected mass flow rate is obtained from the baseline compressor performance map which represents compressor performance with the variable geometry at nominal, scheduled position:

$$w_{\text{base},c,n} = f(N_{c,n}, \Delta p_n). \tag{1}$$

The normalized corrected mass flow rate, which accounts for off-schedule geometry effects, is obtained from the variable-geometry effects performance map

$$w_{\text{var},c,n} = f(CVGP, \Delta P_n). \quad (2)$$

The mass flow rate in the compressor is

$$w = w_{\text{base},c,n}(1 + w_{\text{var},c,n})\left[w_d \sqrt{\frac{T_{in,d}}{T_{in}}} \right]\left(\frac{p_{in}}{p_{in,d}} \right). \quad (3)$$

The normalized adiabatic efficiency value is obtained from the baseline compressor performance map. The adiabatic efficiency, η, is then computed by multiplying the normalized adiabatic efficiency by the design adiabatic efficiency value.

The stagnation temperature rise across the compressor is calculated using isentropic relationships based on an average temperature which is computed using a temperature interpolation constant. The average temperature value is used to determine the constant specific heat of the compressor from curve fit data. The isentropic stagnation temperature rise is then computed and used to find the stagnation temperature at the compressor outlet

$$T_{\text{out}} = \left[\frac{(\Delta T / T_{in})_{\text{ideal}}}{\eta} + 1 \right] T_{in}. \quad (4)$$

The enthalpy corresponding to the temperature at the compressor outlet, h_{out}, is determined from curve fit data based on values of stagnation temperature.

The compressor functions by transmitting mechanical energy (supplied by the shaft) into kinetic energy in the fluid flow. The power or rate of energy transmitted from the shaft to the fluid is

$$P_{\text{comp}} = w(h_{\text{out}} - h_{in}). \quad (5)$$

5.2 Compressor Implementation. In this section we highlight the syntax and input structures that were used to construct the Compressor EngElement, its Domain Model, Ports, and Connector classes. Note that the following excerpts do not constitute a complete Compressor definition.

The Compressor class was defined as a subclass of the DefaultEngElement class (see Fig. 3). This definition appears as

```
public class Compressor
  extends DefaultEngElement {
```

The Compressor class defines a single constructor which is called by the Java Virtual Machine when the class is instantiated. The constructor sets the default name of this compressor instance, and then creates the Compressor's Port objects. The constructor looks as follows:

```
public Compressor () {
  setName(''Compressor'');
  inport = new ZeroDFluidPort(''fluid in'');
  outport = new ZeroDFluidPort(''fluid out'');
  shaftport = new ZeroDMechPort(''shaft in'');
  addPort(inport);
  addPort(outport);
  addPort(shaftport);
  model = new CompressorModel(inport, outport, shaftport);
}
```

The Compressor has three Ports objects: two instances of ZeroDFluidPort and one instance of ZeroDMechPort. The ZeroDFluidPort objects represent the control surface through which the engine working fluid passes. The ZeroDMechPort object defines the surface through which mechanical energy passes (i.e., the mechanical connection with the shaft that drives the compressor blades). Once the Ports are created, they are added to the `ports` Vector defined in DefaultEngElement. Ports must be added to the `ports` Vector in order to correctly build their corresponding graphical Port objects in the Visual Assembly Framework (described in Part 2 of this paper).

The final action taken in the Compressor class constructor is to create an instance of the Compressor's Domain Model class. In this example, the constructor creates a CompressorModel object and references it to the `model` variable. The CompressorModel constructor takes the three Port objects (`inport`, `outport`, and `shaftport`) as arguments.

5.3 Compressor Domain Model. The Compressor's Domain Model is defined by the CompressorModel class, which implements the Model interface (see Fig. 4):

```
public class CompressorModel
  implements Model {
```

The CompressorModel class declares a single constructor, which creates two instances of class DataTable: `fluidDataTable` and `mechDataTable`. These objects are used to define the Port data structure, and thus are passed to each of the Compressor's Port objects.

```
public CompressorModel (ZeroDFluidPort inport, ZeroDFluidPort outport, ZeroDMechPort shaft-
port) {
  fluidDataTable = new DataTable();
  mechDataTable = new DataTable();
  // Both Fluid Ports use same dataTable
  inport.setDataSet(fluidDataTable);
  outport.setDataSet(fluidDataTable);
  // The Shaft Port uses its own dataTable
  shaftport.setDataSet(mechDataTable);
}
```

As stated above, the Compressor defines two instances of ZeroDFluidPort and a single instance of ZeroDMechPort to define the data on its control surfaces. For convenience, the data represented by the Ports is actually stored in a delegate object, called `dataTable`, which is an instance of the DataTable class.

```
public class ZeroDFluidPort
  extends DefaultPort {
  private DataTable dataTable = null;
  public void setDataSet(Object o) {
    table = (DataTable)o;
}
```

The DataTable class provides a structure for storing the data in an instance of the `java.util.Hashtable` class. A Hashtable is a growable list which maps *keys* to *values*. Hashtable stores its values as instances of `java.lang.Object`, which precludes storing primitive data objects such as doubles, unless they are wrapped in objects (e.g., `java.lang.Double`). The DataTable handles this data transformation and wrapping automatically for the developer. The `dataTable` variable is set by calling the `setDataSet` method.

A single instance of DataTable (called `fluidDataTable`) is used to define the data at both the inlet and exit control surfaces of the Compressor. It should be noted that this is not the typical case. More often, separate Port objects will be defined for each control surface. However, for a simple model such as this, we use variable names that indicate whether the parameter is defined at the inlet or exit (e.g., `enthalpyOut` is the computed stagnation enthalpy at the compressor exit).

The data stored in `fluidDataTable` is defined by the Compressor Domain Model. It is defined using get and put methods. The get method in class ZeroDFluidPort is used to get the DataTable (i.e., `dataTable`) from the Port at the other end of a Connector (see Fig. 6), and then invoke the get method on the DataTable.

```
public double get(String s) throws PortConnectionException {
  DataTable t = (DataTable) connector.getDataSet(this);
  return (t.get(s));
}
```

Data computed by the Compressor Domain Model can be defined at its Ports by ''putting'' the data to the Port's DataTable. Similarly, data can be obtained from an adjacent EngElement by ''getting'' the data. The following code excerpt illustrates this:

```
double tind = inport.get(''design temperature'');
double pind = inport.get(''design pressure'');
double toutd = outport.get(''design temperature'');
double poutd = outport.get(''design pressure'');
```

Here the design values of temperature and pressure are obtained from the EngElements connected to the Compressor at its inlet (`inport`) and exit (`outport`). Data can be put on the Compressor's `inport` and `outport` as follows:

```
fluidDataTable.put(''design enthalpy'', designEnthalpyOut);
fluidDataTable.put(''design fuel air ratio'', farind);
```

In this excerpt, the Compressor's design values of enthalpy and fuel air ratio are set at its Ports.

The ZeroDMechPort class shares the same internal structure as ZeroDFluidPort, and is used in the same manner. However, the data being represented will be structural in nature.

5.4 Implementing the Domain Model Behavior. The CompressorModel class defines several methods to implement the mathematical models identified above in Eqs. (1)–(5). The most important of these is the `execute` method, which is called by the Compressor class to compute the operating characteristics of the Compressor:

```
public void execute {
```

The `execute` method first gets the data from the EngElements connected at its inlet and exit fluid connections (i.e., its `inport` and `outport`) and from the Shaft connected to its `shaftport`.

```
  try {
      double tin = inport.get(''temperature'');
      double tind = inport.get(''design temperature'');
      double pin = inport.get(''pressure'');
      double pind = inport.get(''design pressure'');
      double pout = outport.get(''pressure'');
      double tout = outport.get(''temperature'');
      double poutd = outport.get(''design pressure'');
      double hin = inport.get(''enthalpy'');
      double farin = inport.get(''fuel air ratio'');
      double spoolRpm = shaftport.get(''spool rpm'');
      double spoolRpmd = shaftport.get(''design spool rpm'');
```

These data are then used to obtain the normalized inlet-corrected mass flow rate (`correctedNormalizedMdot`) and the normalized adiabatic efficiency value (`normalizedEfficiency`) values from the baseline compressor performance map:

```
double xcom = (pout/pin)/(poutd/pind);
double ycom = (spoolRpm/Math.sqrt(tin))/(spoolRpmd/Math.sqrt(tind));
double[] mdotEtaArray = basePerfMap.
   doubleInterpolationAt(xcom, ycom);// Eq. 1
double correctedNormalizedMdot = mdotEtaArray[0];
double normalizedEfficiency = mdotEtaArray[1];
```

The Compressor's performance maps are represented by instances of the Map class. This class defines a multivariate data structure and methods to read, interpolate and scale the data. In the above code, a double interpolation is performed on the baseline performance map (`basePerfMap`) to return an array containing values for normalized inlet-corrected mass flow rate and normalized adiabatic efficiency at the point (`xcom`, `ycom`).

The adiabatic efficiency is computed by multiplying the normalized adiabatic efficiency by the design adiabatic efficiency value:

```
efficiency = normalizedEfficiency * designEfficiency;
```

The normalized corrected mass flow rate, which accounts for off-schedule geometry effects, is obtained from the variable-geometry effects performance map. This value is time-dependent, so the current simulation time is also obtained:

```
double operatingVariableGeomAngle =
   variableGeomAngleTransientController.
   computeCurrent Value(currentTime);
double variableGeomAngle = biasVariableGeomAngle−
   operatingVariableGeomAngle;

// Determine variable geometry effects value
double[] cShiftArray = correctionPerfMap.
doubleInterpolationAt(variableGeomAngle,ycom);//Eq. 2
double cShift = cShiftArray[0];
```

The Compressor class allows the user to create a control schedule to define the time-dependent value of the variable geometry angle. Schedule data is stored in an instance of TransientControllerSchedule class which is used by an instance of the TransientController class (`variableGeomAngleTransientController`) to obtain the current geometry angle based on the current time (`currentTime`).

The physical mass flow rate (`mdotIn`) in the compressor is:

```
double mdotIn = (correctedNormalizedMdot * (1.0 + cShift))
  *(pin/Math.sqrt(tin))*(designMdotIn *
  Math.sqrt(tind)/pind)*weightFlowCorrectionCoef;//Eq. 3
```

The stagnation temperature rise across the compressor (`tempRiseParamIdeal`) is calculated using isentropic relationships based on an average temperature which is computed using a temperature interpolation constant. The average temperature value is used to determine the constant specific heat of the compressor from curve fit data. The isentropic stagnation temperature rise is then computed and used to find the stagnation temperature (`toutp`) at the compressor outlet:

```
double pratio = pout/pin;
double tempRiseParamIdeal = this.tempRiseParamIdeal(tin, tout, farin, pratio);
double toutp = tin * (1.0+(tempRiseParamIdea1/efficiency * temperatureCorrectionCoef)));//Eq.
4
```

The enthalpy corresponding to the temperature at the compressor outlet (`enthalpyOut`) is determined from curve fit data based on values of stagnation temperature (`toutp`) and the fuel-air ratio (f/a) in the compressor (`farin`).

```
double enthalpyOut = GasStates.computeGasEnthalpyWith(toutp, farin);
```

The power transmitted from the shaft to the fluid is computed as:

```
double energyTerm = (enthalpyOut−hin)*mdotIn;//Eq. 5
```

Finally, the newly computed values are used to define the values at the Ports so that connected EngElements can access that data:

```
fluidDataTable.put(''enthalpy'',enthalpyOut);
mechDataTable.put(''mech energy term'',energyTerm);
fluidDataTable.put(''fuel air ratio'',farin);
fluidDataTable.put(''mass flow rate'',mdotIn);
}
```

5.5 Connectors. Two connector classes were defined for the example: ZeroDFluidConnector and ZeroDMechanicalConnector. The ZeroDFluidConnector connects two ZeroDFluidPorts and the ZeroDMechanicalConnector connects two ZeroDMechanicalPorts. Both are subclasses of DefaultConnector (see Fig. 6). Because both types of connections are homogeneous (i.e., single discipline, single level of fidelity), there is no need for a Transform object in either ZeroDFluidConnector or ZeroDMechanicalConnector. The class definition for ZeroDFluidConnector is shown below. The ZeroDMechanicalConnector was similarly defined.

```
public class ZeroDFluidConnector
  extends DefaultConnector {
  public ZeroDFluidConnector(ZeroDFluidPort port1, ZeroDFluidPort port2) {
    super(port1, port2);
    port1.setConnector(this);
    port2.setConnector(this);
  }
}
```

5.6 Creating and Assembling Components. Users can instantiate and assemble components programmatically as shown in the following code (alternatively, the Visual Assembly Framework which is described in Part 2, can be used to graphically assemble components). This code creates the Inlet and Fan components for the E^3 model, defines specific data and connects them:

```
EngElement inlet = new Inlet(''INLET'');
inlet.setDesignMassFlowRate(1260.3);
inlet.setPressureRecovery(0.988);
...
EngElement fan = new Compressor(''FAN'');
fan.setDesignMassFlowRate(1260.3); //(1bm/sec)
fan.setDesignEfficiency(0.9678); //[dimensionless]
...
Port fanInFluidPort = fan.getPort(''fluid in'');
Port inletOutFluidPort=inlet.getPort(''fluid out'');
Connector c1 = new Connector(inletOutFluidPort,fanInFluidPort);
```

6 Conclusion

Designing and developing new aerospace propulsion technologies is a time-consuming and expensive process. Computational simulation is a promising means for alleviating this cost, due to the flexibility it provides for rapid and relatively inexpensive evaluation of alternative designs, and because it can be used to integrate multidisciplinary analysis earlier in the design process. However, integrating advanced computational simulation analysis methods such as CFD and FEA into a computational simulation software system is a challenge. A prerequisite for the successful implementation of such a program is the development of an effective simulation framework for the representation of engine components, subcomponents, and subassemblies. To promote concurrent engineering, the framework must be capable of housing multiple views of each component, including those views which may be of different fidelity or discipline. Object-oriented technology improves design and development of aerospace simulation systems by leveraging proven software design to produce a reusable component-based architecture which can be extended and customized to meet future application requirements.

A common engineering model formalism for representing gas-turbine engine systems, subsystems, and components was developed based on hierarchical decomposition. This approach provides a simple and intuitive method to compose more complex engine component models out of simpler ones. The composition process was made possible by establishing well-defined class interfaces—EngElement, Domain Model, Port, and Connector.

Robust simulation requires representing models at different levels of abstraction. Such multimodels promote concurrent engineering by housing multiple views of the component, including those views which may be of different fidelity or discipline. The Common Engineering Model design supports this through the use of Domain Model classes which encapsulate the abstractions and remove them from the EngElement structure. This was done to promote flexibility, allowing EngElement and Domain Models to vary independently. A consequence of this design is that it encourages the wrapping of existing software codes as domain models, making possible the integration of existing analysis methods into the simulation.

Component interaction was defined by Connector and Port classes. Structure of the Port classes is tied to the discipline and fidelity level of the component's Domain Model. Disciplinary and fidelity mismatches between engine components are handled by Transform objects in a Connector, allowing for ''zooming'' capabilities and mapping between discipline domains.

The Common Engineering Model provides the foundation for the development of component-based models for the simulation of gas turbine engines in Onyx. Due to their standardized interfaces, EngElements may be represented visually, distributed on remote machines, or downloaded from Web servers and plugged into Onyx. These uses are discussed in Part 2 of this paper.

Acknowledgments

The work described in this paper was made possible by funding from the NASA Glenn Research Center Computing and Interdisciplinary Systems Office, and the University of Toledo. We would like to thank Greg Follen at NASA Glenn for his continued support.

Nomenclature

CVGP = compressor variable geometry position
h = specific enthalpy
N = angular velocity
p = pressure
P = power
T = temperature
t = time
w = mass flow rate
η = compressor adiabatic efficiency

Subscripts

base = baseline
c = corrected
comp = compressor
d = design point
ideal = isentropic
in = inlet
n = normalized
out = outlet
var = variable

References

[1] Evans, A., Lytle, J., Follen, G., and Lopez, I., 1997, ''An Integrated Computing and Interdisciplinary Systems Approach to Aeropropulsion Simulation,'' ASME Paper 97-GT-303.
[2] Jameson, A., 1997, ''Re-Engineering the Design Process Through Computation,'' AIAA Paper No. 97–0641.
[3] Fawke, A. J., Saravanamuttoo, H. I. H., and Holmes, M., 1972, ''Experimental Verification of a Digital Computer Simulation Method for Predicting Gas turbine Dynamic Behavior,'' The Institution of Mechanical Engineers Combustion Engines Group. Vol. 186, p. 32.
[4] Koenig, R. W., and Fishbach, L. H., 1972, ''GENENG A Program for Calculating Design and Off-Design Performance of Turbojet and Turbofan Engines,'' NASA TN D-6552.
[5] Seldner, K., Mihailowe, J. R., and Blaha, R. J., 1972, ''Generalized Simulation Technique for Turbojet Engine System Analysis,'' NASA TN D-6610.
[6] Seller, J., and Daniele, C. J., 1975, ''DYGEN A Program for Calculating Steady-State and Transient Performance of Turbojet and Turbofan Engines,'' NASA TND-7901.
[7] Drummond, C., Follen, G., and Cannon, M., 1994, ''Object-Oriented Technology for Compressor Simulation,'' AIAA Paper 94–3095.
[8] Claus, R. W., Evans, A. L., Lytle, J. K., and Nichols, L. D., 1991, ''Numerical Propulsion System Simulation,'' Comput. Syst. Eng., **2**, No. 4, pp. 357–364.
[9] Arnold, K., and Gosling, J., 1996, *The Java Programming Language*, Addison Wesley, Reading, MA.
[10] Fishwick, P. A., 1997, ''Computer Simulation: Growth Through Extension,'' Transactions of The SCS, **14**, pp. 13–23.

[11] Booch, C., 1991, *Object Oriented Design with Applications*, Benjamin/Cummings, New York, NY.
[12] Snyder, A., 1986, ''Encapsulation and Inheritance in Object-Oriented Languages,'' in *Object-Oriented Programming, Systems. Languages, and Applications Conference Proceedings*, Portland, OR, pp. 38–45.
[13] Fishwick, P. A., and Zeigler, B. P., 1992, ''A Multimodel Methodology for Qualitative Model Engineering,'' ACM Trans. Mode. Comput. Simu., **12**, pp. 52–81.
[14] Szuch, J. R., 1974, ''HYDES—A Generalized Hybrid Computer Program for Studying Turbojet or Turbofan Engine Dynamics,'' NASA TM X-3014.
[15] Schobeiri, T., 1986, ''A General Computational Method for Simulation and Prediction of Transient Behavior of Gas Turbines,'' ASME Paper 86-GT-180.
[16] Stewart, M., 1995, ''Axisymmetric Aerodynamic Numerical Analysis of a Turbofan Engine,'' ASME Paper No. 95-GT-338.
[17] Schlichting, H., 1979, *Boundary-Layer Theory*, 7th ed., McGraw-Hill, New York, NY.
[18] Gamma, E., Helm, R. Johnson, R., and Vlissides, J., 1995, *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison Wesley, Inc., Reading, MA.
[19] Hall, E. J., and Delaney, R. A., 1995, ''Investigation of Advanced Counterrotation Blade Configuration Concepts for High Speed Turboprop Systems: Task VII-ADPAC Users Manual,'' NASA CR 195472.
[20] Claus, R. W., Evans, A. L., and Follen, G. J., 1992, ''Multidisciplinary Propulsion Simulation using NPSS,'' AIAA Paper 92-4709.
[21] Holt, G., and Phillips, R., 1991, ''Object-Oriented Programming in NPSS Phase II Report,'' NASA CR-NAS3-25951.
[22] Davis, D. Y., and Steams, E. M., 1985, ''Energy Efficient Engine—Flight Propulsion System Final Design and Analysis,'' NASA CR-168219.
[23] Daniele, C. J., Krosel, S. M., Szuch, J. R., and Westerkamp, E. J., 1983, ''Digital Computer Program for Generating Dynamic Engine Models (DIGTEM),'' NASA TM-83446.
[24] Reed, J. A., 1998, ''Onyx: An Object-Oriented Framework for Computational Simulation of Gas Turbine Systems,'' Ph.D. dissertation, The University of Toledo, Toledo, OH.

Computational Simulation of Gas Turbines: Part 2—Extensible Domain Framework

John A. Reed
e-mail: jreed@memslab.eng.utoledo.edu

Abdollah A. Afjeh

Mechanical, Industrial and Manufacturing Engineering Department,
The University of Toledo,
Toledo, OH 43606

This paper describes the design concepts and object-oriented architecture of Onyx, an extensible domain framework for computational simulation of gas turbine engines. Onyx provides a flexible environment for defining, modifying, and simulating the component-based gas turbine models described in Part 1 of this paper. Using advanced object-oriented technologies such as design patterns and frameworks, Onyx enables users to customize and extend the framework to add new functionality or adapt simulation behavior as required. A customizable visual interface provides high-level symbolic control of propulsion system construction and execution. For computationally-intensive analysis, components may be distributed across heterogeneous computing architectures and operating systems. A distributed gas turbine engine model is developed and simulated to illustrate the use of the framework. [S0742-4795(00)02403-0]

Keywords: Object-Oriented Modeling, Gas Turbines, Computational Simulation

1 Introduction

The design of gas turbine engine simulation software has traditionally focused on developing single, custom-built systems. As a result, leveraging software design has generally been done in an ad hoc manner. However, it is widely recognized that such an approach, while adequate for simple systems, is not cost effective for realizing large, complex systems. Systematic software reuse—the process of creating software systems from predefined software components—offers a methodology for moving from the practice of hand-crafted systems to a more formal approach based on engineering principles [1]. Object-oriented technology (OOT) has been embraced by many as a sufficient means for software reuse, but OOT does not, by itself, promote large-scale system reuse.

Domain analysis and modeling is an emerging collection of software design methodologies which aim to identify the essential features, components, capabilities, interfaces, abstractions, etc., of a family of systems in a domain [2]. Based mostly on object-oriented techniques, they enhance those techniques with respect to design for reusability. In recent years, two approaches to domain modeling have emerged: integrative modeling and generative modeling. Integrative modeling directly extends the object-oriented modeling methodologies and typically defines an integrated set of submodels in the form of object composition/aggregation and generalization/specialization hierarchies. Generative models primarily deal with software architecture issues such as identifying the fundamental programming abstractions in the domain, creating libraries of interoperable software components, and defining knowledge-representation languages for combining and coupling software ''building blocks.'' Two special cases of generative models—*design patterns* and *object-oriented frameworks*—have been shown to be beneficial to the development of reusable and flexible, domain-specific software systems.

Design patterns [3] are a promising technique for achieving widespread reuse of software architectures. A design pattern is a recurring solution to problems that arise when building software in various domains. Patterns aid the development of reusable components and frameworks by expressing the structure and collaboration of participants in a software architecture at a level higher than (i) source code or (ii) object-oriented design models that focus on individual objects and classes [4]. Patterns also are particularly useful for documenting software architectures and design abstractions. They provide a common and concise vocabulary which is useful in conveying the purpose of a given software design.

While patterns are useful in promoting reuse, they are not sufficient to create flexible software by themselves. Patterns are, in general, language-independent and do not directly produce reusable code [5]. Therefore, it is important to apply patterns in a manner which capitalizes on the reuse of abstract design, architecture knowledge and code. Object-oriented frameworks do precisely this.

An object-oriented framework is a set of classes that embodies an abstract design for solutions to a family of related problems [6]. The set of classes define ''semi-complete'' applications that capture domain-specific object structures and functionality. Specific functionality in new applications is realized by inheriting from, or composing with, framework components.

A major product of domain modeling is the identification of software components—self contained software elements which can be controlled dynamically and assembled to form applications [7]. The central step in identifying them is recognizing recurring fundamental abstractions in the domain. By identifying these abstractions and standardizing their interfaces, these components become interchangeable. Such components are said to be ''plug-compatible'' as they permit components to be ''plugged'' into frameworks without redesign.

This paper describes the architecture of *Onyx*, an object-oriented domain framework for the modeling and simulation of gas turbine systems. Onyx leverages object-oriented design techniques—design patterns, frameworks, and software components—to provide a flexible and extensible environment which can be used to compose new engine component models, to inspect and edit existing models, and simulate and display execution results. These models, which are based on the Common Engineering Model described in Part 1 of this paper, provide plug-compatible software components which users can combine to form increasingly complex gas turbine models. The following section provides an overview of the structure of the Onyx framework and the Java™ programming platform [8], which is used by Onyx.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-347. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

The next three sections describe the major features of the framework, including the graphical user interface, distributed services, and execution control capabilities. The final section illustrates how Onyx is used to build and simulate gas turbine engine models. An engine model is visually composed using the example engine model introduced in Part 1, and a distributed simulation is demonstrated.

2 Onyx

Onyx is arranged as a layered collection of four individual frameworks:

1 ***Engine Component Framework (ECF)***. This framework contains collections of software components which represent physical components in the gas turbine domain (e.g., compressor, turbine, combustor, etc.). The components are based on the Common Engineering Model defined in Part 1 and are represented in the ECF by EngElement and Port objects, which are connected by Connector objects.

2 ***Visual Assembly Framework (VAF)***. This framework ''sits'' on top of the Engine Component Framework and provides visual representations of the objects in the Engine Component Framework. The VAF forms a major part of the Onyx Graphical User Interface (GUI) and provides tools to visually assemble and manipulate simulation models in the ECF.

3 ***Connection Services Framework (CSF)***. The CSF is an extensible mechanism which provides both local and distributed connection services for the ECF and VAF. The CSF is responsible for locating and loading both local and distributed components, and making them available for instantiation in the ECF and VAF.

4 ***Execution Control Framework (XCF)***. This framework provides a set of tools with which to control Onyx simulations. The XCF defines a high-level abstraction mechanism for creating execution strategies for the simulation, as well as registration and selection of numerical solvers for solution of the system of equations.

This layering of frameworks enhances Onyx's flexibility as the individual frameworks are loosely coupled. This allows each framework to be used more-or-less independently. For example, the Engine Component Framework, Connection Services and Execution Control Frameworks can be used without the Visual Assembly Framework. From a design perspective, the reduction of coupling between frameworks is beneficial as it helps to minimize the effect that a design change to one framework has on another.

3 Visual Assembly Framework

Engineers often use schematic drawings to represent gas turbine systems and subsystems. It is natural then to represent computational simulations of such entities using this visual metaphor. The Visual Assembly Framework (VAF) provides a high-level architecture for visually creating, assembling, and manipulating engine schematics based on engine component models developed using EngElement classes.

The design requirements of the Visual Assembly Framework were the following: (i) visual elements are needed to represent the objects which form Onyx's engine component model; (ii) the concept of component composition developed in Part 1 must also be supported visually; and (iii) the framework must take care of managing basic graphical functions—window management, displaying objects, moving and dragging visual elements, tracking mouse movements, etc. This reduces the programming burden for developers using the framework.

In addition to these goals were some constraints. First, the framework should decouple the visual user interface (UI) objects from their counterparts in the Engine Component Framework. Such an approach would allow a component's UI to be changed easily, possibly at run-time without affecting the structure of the engine component object. Second, the design should allow the developer to override the default visual representations as much as is practically possible.

The Java platform was selected for developing Onyx in part due to its integrated graphical support. Onyx uses the Swing component set—a subset of the Java Foundation Classes (JFC) [9]—to implement its graphic interface. Because Swing component are lightweight (i.e., not based on the native windowing system), they are very flexible and have the ability to support multiple look-and-feel standards (e.g., Windows™, Motif™ and Macintosh™).

3.1 Overview of the Visual Assembly Framework. A gas turbine simulation model is developed by building a schematic diagram of the gas turbine in a SchematicFrame window (see Fig. 1). Individual engine components, represented graphically as SchematicIcons, are selected from the ToolBox window and dragged into a SchematicFrame. The SchematicIcons are then interconnected to form a schematic representation.

A SchematicIcon is composed of a VEngElement and one or more VPorts. VEngElement is the visual analog of the EngElement class in the Engine Component Framework (ECF). By default, it is composed of an Icon, which presents an image of the engine component, and a Label, which displays the name of the EngElement object it represents. VPorts are the visual representation of the Port objects in the ECF. VPorts are the connection points between VEngElements, and are color-coded to represent the Port's type (e.g., fluid transfer, mechanical, etc.). VConnectors (i.e., *visual* Connectors) define a graphical connection between two VPorts, and represent the Connector objects in the ECF.

Hierarchical engine schematic structures can be created in the VAF using SchematicIcons and SchematicFrames. The approach is based in-part on the work of Curlett et al. [10]. A hierarchical model is graphically constructed by adding a Composite SchematicIcon to a model in an existing SchematicFrame. This creates a new SchematicFrame which represents a deeper (more refined) level in the structural hierarchy. At the same time, it creates a CompositeEngElement in the ECF; this is represented by the SchematicFrame. Additional SchematicIcons can then be dragged into this new SchematicFrame to form a sub-model schematic. In Fig. 1, the SchematicIcon labeled Core is a CompositeEngElement. Each of the EngElements in the Core composite are also represented by SchematicIcons. These are seen as a separate schematic in the SchematicFrame window titled Core.

SchematicIcons in the Core sub-model are connected to SchematicIcons in the parent model by *promoting* the sub-model SchematicIcon's VPort. Every SchematicFrame in a simulation (except the top-level SchematicFrame), has a *promotion border*, which appears as a solid border around the inner edge of the SchematicFrame. When a connection is made from a VPort into the promotion border, the VPort is connected to a BorderVPort in the promotion border, and a new VPort is added to the Composite SchematicIcon in the parent model. For example, in Fig. 1, the Compressor SchematicIcon's input VPort (on the left-hand side of the SchematicIcon) has been promoted. The VPort is depicted in the Turbofan SchematicFrame as a VPort on the left-hand side of the Composite SchematicIcon. This VPort may now be connected to any compatible VPort in the parent model schematic. In this case, it is connected to the MV13 SchematicIcon. Similarly, the HP Turbine SchematicIcon's exit VPort has been promoted and appears on the right-hand side of the Composite Schematic Icon. It is connected to the MV41 SchematicIcon in the parent model schematic.

The location of a promoted VPort on the Composite SchematicIcon corresponds to the relative location of the BorderVPort location in the promotion border. A BorderVPort can be dragged to any point within the promotion border, and the corresponding VPort is moved accordingly.

Each VEngElement, VPort and VConnector has a popup menu associated with it. The menu allows the user to access various functions such as moving, deleting, copying, etc. In the VEngEle-

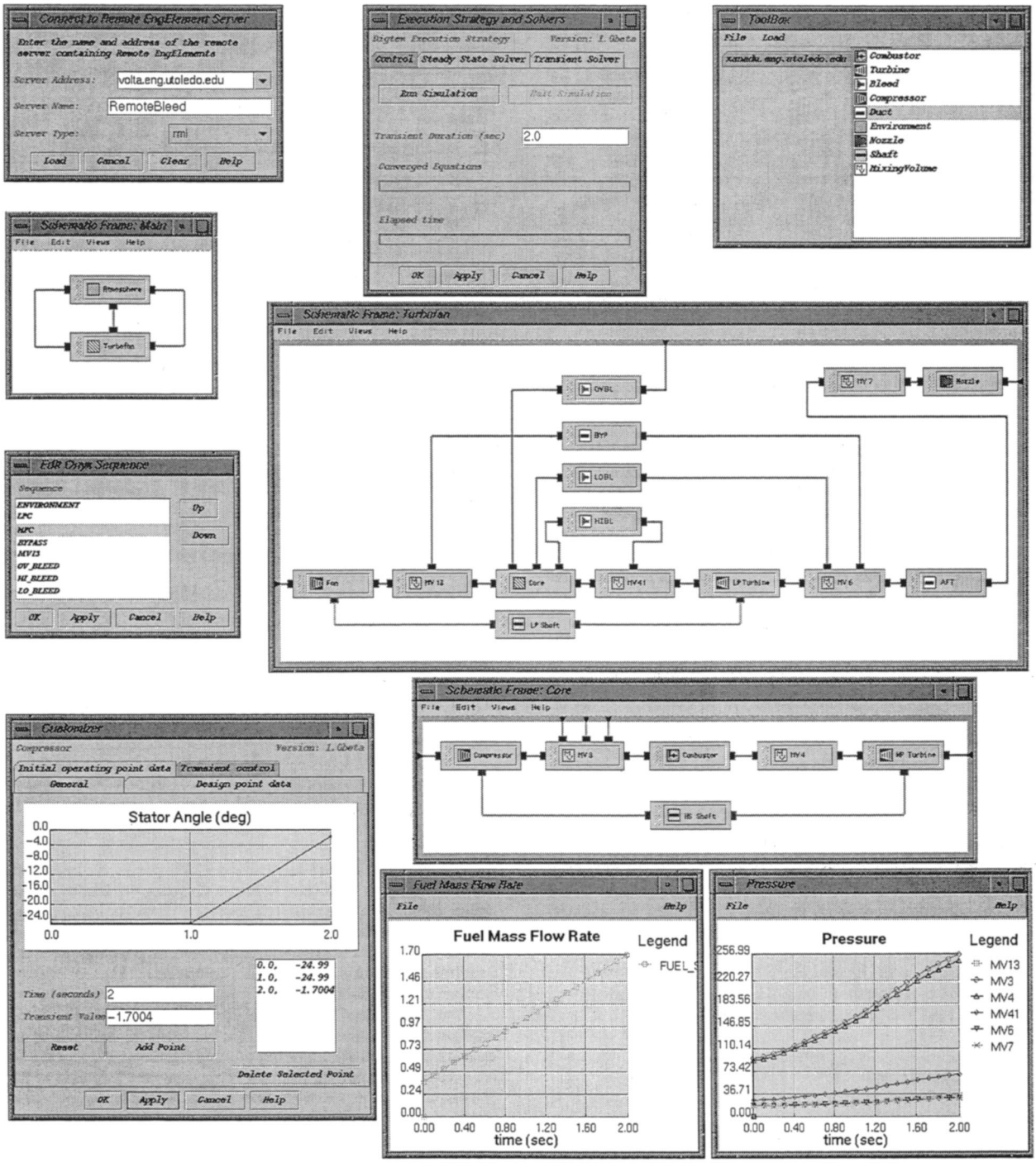


Fig. 1 Overview of Onyx Visual Assembly Framework

ment, the popup menu has a special item for ''customizing'' the component. When selected, a Customizer object is displayed. Customizers are graphical interfaces which allow the user to change an EngElement's attributes (see lower-left corner of Fig. 1). Typically, these are used to supply user-defined data to the EngElement's Domain Model. The VAF also provides additional tools for plotting, editing files, and displaying Web browsers to search on-line documentation.

3.2 SchematicIcons. An objective of the Visual Assembly Framework (VAF) is to provide a SchematicIcon for each type of EngElement available in the Engine Component Framework (ECF). Each EngElement type may have different numbers and types of Ports, as well as different Domain Models. Consequently, the VAF must provide SchematicIcons with VPorts and Customizers that match the EngElement's type. One way to account for this variability in the VAF design is to represent SchematicIcon as an abstract class. Subclasses can then define appropriate Icons, Customizers, and VPorts for a particular EngElement type. However, this approach leads to a very broad and shallow inheritance tree, indicating poor use of inheritance.

A more flexible approach is to create SchematicIcons using object composition. The VAF uses a variation of the JavaBeans™ ''BeanInfo'' class approach [7] to compose SchematicIcons with the correct attributes (i.e., Icon, Customizer, and VPorts). In this

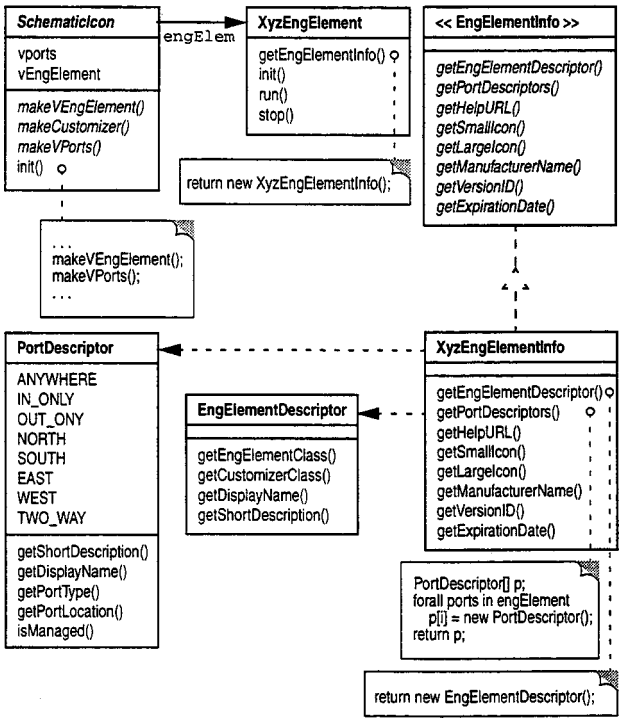


Fig. 2 SchematicIcon ''Info class'' structure diagram

technique, each EngElement is provided with a separate auxiliary ''Info class'' which identifies and encapsulates its UI-specific attributes. The benefit of using an Info class is that it separates the UI part of the engine component representation from the underlying component model defined by the EngElement. This is extremely useful in reducing the complexity of the EngElement. It also provides flexibility, since the EngElement's UI may be redefined (at runtime, if desired) without affecting the EngElement itself.

The ''Info class'' mechanism centers around the use of the EngElementInfo interface (see Fig. 2). The interface defines a set of methods for obtaining UI information for an EngElement. For example, the `getSmallIcon` method returns a reference to the Icon used in the SchematicIcon. The interface also defines methods for obtaining information for use in objects other than the SchematicIcon. The `getHelpURL` method is used by the Onyx help system to obtain the Uniform Resource Locator (URL) pointing to the location of the EngElement's help file.

The `getEngElementDescriptor` and `getPortDescriptors` methods in the EngElementInfo interface are particularly important. The `getEngElementDescriptor` method returns an EngElementDescriptor object which encapsulates information about the EngElement. This information includes the EngElement class name, display name, a short description of its function, and the class name of the Customizer used to customize its Domain Model. The method `getPortDescriptor` returns information concerning the number, type, display name and location of the VPorts to be included in the SchematicIcon. This information reflects the number, type, and names of the Port objects in the EngElement.

The EngElementInfo class is implemented by other classes. In the figure, the XyzEngElementInfo class implements the interface and defines concrete methods to return information about the XyzEngElementInfo class. XyzEngElement returns its EngElementInfo class when its `getEngElementInfo` method is invoked.

The abstract SchematicIcon class' `init` method defines a series of steps for creating a SchematicIcon to represent an EngElement object. It obtains information about the EngElement object's UI by getting the ''Info class'' objects, then uses that information to create UI-specific versions of the VEngElement and VPorts, which are then added to the SchematicIcon to define its appearance.

3.3 Customizer. A customizer is a UI element which allows the simulation user to edit an EngElement's attributes. These attributes include general data, such as the EngElement name, class name, and description, as well as the data contained in Domain Model. A natural consequence of this relationship is that there is a tight coupling between the two classes: the structure of the customizer is dictated by the data structure of the EngElement. This posed a challenge in developing a customizer class structure. As a graphical component, a customizer should be defined as part of the VAF. However, attempting to create a ''universal'' customizer for all types of EngElements would be very difficult. It is more practical to let the developer of the EngElement create the customizer, than it is to define its structure as part of the VAF. To successfully support this approach, the design must address two key issues: (i) customizers which are created by the developer must be *pluggable* into the VAF; that is, they must be easily integrated into the framework; and (ii) developers should be able to create customizers as easily as possible.

In order to be pluggable into the VAF as a customizer, a class must implement the Customizer interface (see Fig. 3). This Java interface defines two methods. The first method, `setTarget`, is invoked after Customizer construction to identify the EngElement object that it is being customized. The `commitChanges` method is called by the VAF when the user accepts any changes made in the customizer's data fields. Besides implementing these methods, a customizer must be a subclass of `java.awt.Component`. This requirement is necessary so that it can be added to an instance of CustomizerFrame.

To maximize ease of use, the VAF allows developers to subclass the CustomizerPage class, compose it with desired UI objects, and add it to BasicTabbedCustomizer. CustomizerPage provides methods to handle common issues such as laying out components. Since BasicTabbedCustomizer adds instances of Customizer, it is also possible to add classes which inherit from `java.awt.Component` and implement Customizer. Figure 1 depicts an instance of BasicTabbedCustomizer, with instances of several CustomizerPage classes added. These pages, titled *General*, *Design point data*, *Initial operating point data*, and *Transient control*, contain specific groupings of data fields and other UI widgets.

The Customizer class structure provides considerable flexibility. It allows the user to compose the UI or inherit functionality

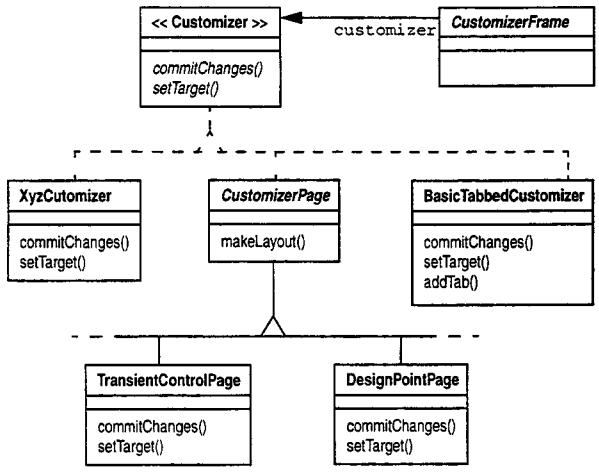


Fig. 3 Customizer class structure diagram

and structure when developing a customizer. By adhering to an interface, users can develop different customizers and plug them into the VAF as desired. Furthermore, a customizer is a subclasses of `java.awt.Component`, so users can use Java Integrated Development Environments (IDEs) to quickly construct customizers from AWT or Swing JavaBean GUI components.

3.4 Organizing Class Information. The design of the VAF emphasizes the distribution of behavior among different classes. An EngElement's behavior is delegated to its Domain Model; its UI is spread out among the EngElementInfo, PortDescriptor, EngElementDescriptor and Customizer classes. Onyx uses Java Archive (JAR) files to organize and maintain these collections of files. An EngElement class, its auxiliary classes and any data (such as the UI's Icon image), can be archived in a JAR file, and loaded into Onyx for use in both the VAF and ECF. The JAR file format is also useful for delivering EngElement classes across the Internet, as they can be compressed to reduce file size. Furthermore, JAR files can have a digital signature added to identify the author of the files contained in the JAR.

4 Connection Services Framework

Higher-fidelity, multidisciplinary computational simulation of the gas turbine engine is computationally intensive. It is estimated that a single disciplinary analysis, such as an aerodynamic, unsteady, three-dimensional, viscous simulation of a complete engine, requires on the order of 10^{12} floating point operations per second, with full engine multidisciplinary analysis 2 to 3 orders of magnitude in excess of that estimate [11]. While such resources are not currently available, advanced computing techniques are expected to make integrated interdisciplinary analysis practical in the near future. It is essential, therefore, that future gas turbine propulsion simulation systems be capable of accessing extensive computational resources.

Much of the recent research aimed at increasing computational performance has concentrated on parallel processing, wherein a large computational problem is decomposed into many smaller tasks. Until recently, these efforts have focused on the use of massively parallel computers, which combine thousands of processors and hundreds of gigabytes of memory in a single unit, to give enormous computational power. However, with the emergence of low-cost, high-performance microprocessor-based computers and local and wide-area networks, there is growing movement away from highly centralized structures to decentralized distributed structures. *Distributed parallel computing*, which uses general-purpose workstations connected by a network as a large parallel computing resource, is a promising trend in parallel processing. With this technique, existing computing resources can be utilized effectively to achieve supercomputer-level power without the additional cost of a supercomputer.

The object-oriented paradigm is naturally suited towards distributed computing as object-oriented techniques encourage the decomposition of the problem space into pieces—objects—that collaborate to accomplish more complex tasks. Objects may be defined to perform a specific piece of the computation and placed on different computers; these objects then communicate with one another via messages to carry out the complete computation.

In contrast to *local* objects, which exist in a single address space, *distributed* objects exist in different address spaces. A consequence of this is that the pointers—entities used to identify the objects location in memory—in a local address space are not valid in another (remote) address space. To circumvent this problem, distributed object infrastructures—generally referred to as an *object request broker*, or *ORB*—provide an underlying mechanism to eliminate the boundary between local and remote computing. Using an ORB, an object reference is created locally and bound to a remote object. Methods may be invoked on the reference as if it were a local object. The ORB transparently intercepts these method invocations and transmits the method request and its arguments to the proxy object. The ORB marshals and unmarshals the method arguments and return values for inter-address space communication.

Distributed object infrastructures permit the developer of distributed applications to adopt a unified model of object interaction that supports location, platform, and programming language transparency. Such transparency is not without its costs, however, and over the years, the designers of some distributed object architectures have elected to forego some neutrality in exchange for perceived improvements in performance, applicability to specific tasks, and/or ease of use. As a result, there are number of distributed object infrastructures in use today. The most popular and widely used are: *RMI* (Remote Method Invocation), the Java platform's distributed object infrastructure [12]; *Voyager*, a full-featured Java ORB architecture from ObjectSpace [13]; *CORBA* (Common Object Request Broker Architecture), a platform-neutral standard which provides flexible object communication and activation in distributed heterogeneous object-oriented computing environments [14]; and *DCOM* (Distributed Component Object Model), Microsoft Corporation's Windows-centric distributed object protocol [15].

4.1 Web-based Distribution. Since its inception in 1990, the World Wide Web (WWW or Web) has quickly emerged as a powerful tool for connecting people and information on a global scale. Built on broadly accepted protocols, the WWW removes incompatibilities between computer systems, resulting in an ''explosion of accessibility'' [16]. These protocols have been extended and integrated with other new related technologies that provide for the delivery of content that is much more dynamic in nature. The most important of these related developments has been the introduction and widespread adoption of the Java programming language as a standard for Internet-based computation.

The integration of the Web and Java represents a technological advancement that enables a fundamentally new approach to modeling and simulation. Java's platform and language independence, combined with the Web's efficient and ubiquitous delivery mechanism, facilitate the widespread deployment of computational models. Standardized models, in the form of software components, can be placed on Web servers, downloaded across the Web to client machines, and loaded into simulation systems for use by engineers. Such an approach fosters model reuse, and collaboration in the design process. As a result, web technology has the potential to significantly alter the ways in which simulation models are developed (collaborative, by composition), documented (dynamically, using multimedia), analyzed (open, widespread investigation) and executed (using massive distribution) [17].

4.2 Distributed and Web-based Connection Strategies in Onyx. Onyx's Connection Services Framework (CSF) provides developers with a simple and extensible mechanism for integrating both distributed object computing and web-based modeling strategies into a gas turbine engine simulation. The main objective in designing the CSF was to develop a uniform and consistent mechanism for instantiating and/or referencing EngElements. In order to meet this objective, the CSF must be capable of accommodating the various available distributed object infrastructures which might be utilized (e.g., CORBA, RMI, etc.), as well as incorporating local object instantiation and web-based object downloading. Furthermore, the CSF was designed so that future distribution strategies can be incorporated within the CSF. Figure 4 illustrates how these various strategies are implemented in Onyx. In (i), the *distributed object strategy*, EngElements are placed on remote servers and connected to Onyx by various distributed object infrastructures (e.g., RMI, CORBA, etc.). In (ii), the *web-based distribution strategy*, Java classes, in the form of byte-codes, are downloaded from a Web server to the local file system, loaded into a Java VM and instantiated for use in Onyx. And in (iii), the *local connection strategy*, EngElement class files

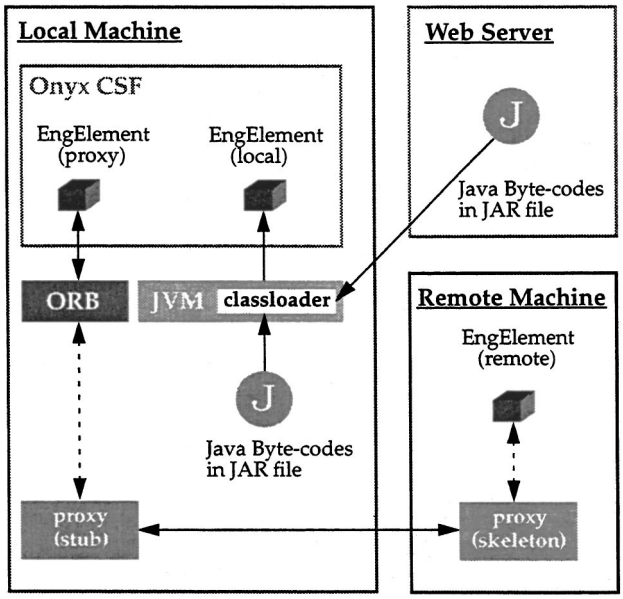


Fig. 4 Overview of the services provided by CSF

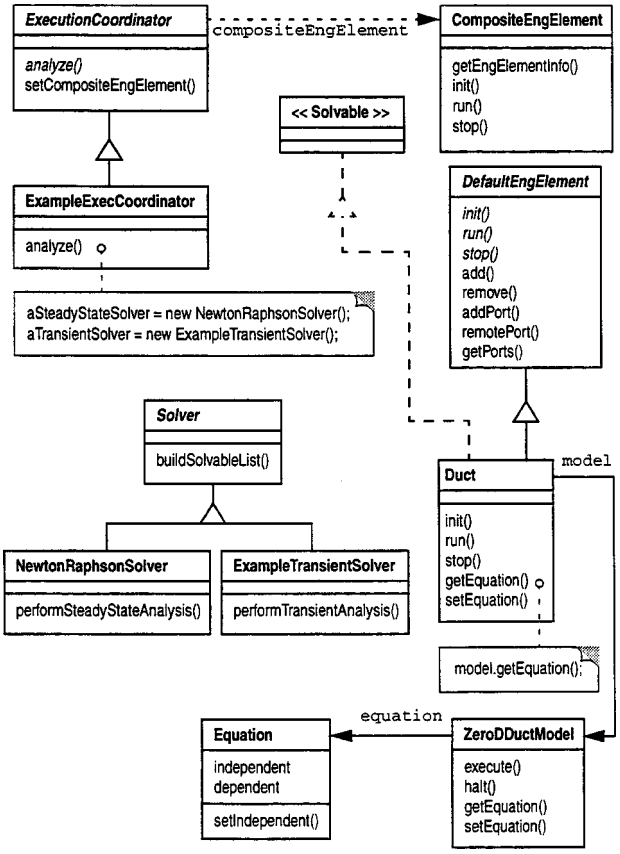


Fig. 5 XCF class structure diagram

are read from the local file system, loaded into a Java VM and instantiated for use in Onyx. This flexible approach offers the following advantages:

- ability to distribute a computationally intensive process across a number of processors
- ability to leverage legacy code limited to platforms offering specific programming and/or operating systems
- facilitation of collaborative exchange of EngElements in the design process
- specialization of computer execution environment (i.e., placement of codes on appropriate computing platforms; such as visualization codes on high-end graphic workstations; computationally intensive codes on supercomputers, etc.)

Connections to distributed and local resources are made by the user from a graphical user interface component known as the ToolBox (see top right corner of Fig. 1). Selections made from ToolBox's Load menu bring up appropriate Dialogs to: load a local EngElement; load a Web-based EngElement; or connect to a distributed EngElement on a remote machine (top left corner of Fig. 1). Once connections are established, Onyx lists the EngElements available from the selected host. In the case of local and distributed servers, the name of the host appears on the right as tabbed panes, while the names of available EngElements appear in the list on the right. For web-based EngElements, the byte-codes are downloaded to the local file system, so the name of the EngElement appears in the list for the locally available EngElements. EngElements are instantiated only when selected by the user from the list for placement into a SchematicFrame. This technique, known as lazy initialization, is used to reduce Onyx's memory footprint.

Onyx currently supports local and web-based EngElement loading, and distributed objects using RMI and Voyager. Support for CORBA will be added in the near future.

5 Execution Control Framework

Model execution is a key step in the process of performing a simulation. In most cases, it involves finding a solution to the mathematical equation(s) which define the behavior of the model. The exact nature of the solution process is dependent on a number of issues, such as the complexity of the problem geometry, the nature of the defining equations, and the imposition of special boundary and initial conditions. In cases where the equations, geometry, and boundary conditions are of general form, the solution can be obtained using standard numerical approaches. However, in complex cases the solution algorithm is frequently tailored to handle the specific characteristics of the model.

This relationship between solution algorithm and model is important in the design of Onyx. As described in Part 1, Onyx's Common Engineering Model allows users to develop heterogeneous simulations based on collections of hierarchically connected models. These models can represent the system at different levels of abstraction, and consequently may have various levels of complexity. It is possible that each different model may require a specific solution algorithm, thus precluding the use of a single, generalized approach for executing the overall simulation. This situation forces us to adopt an approach which allows different solution algorithms to be integrated into the framework.

A solution algorithm primarily defines two entities: a *solver* and a *controlling algorithm*. The solver is the software equivalent of the numerical method which is used to obtain an exact or approximate solution to the equation or sets of equations. The controlling algorithm provides support for the solver, through actions such as setting boundary conditions, updating counters, streaming results to a file, etc. In most simulation software, there is no distinction made between the two; the solution algorithm *is* the solver.

In heterogeneous simulations, it is advantageous to provide separation between solver and controlling algorithm. This allows for the construction of a single controlling algorithm for the complete simulation. Such an approach can be essential to carry out complex simulations in which the objective functions of the individual solvers are competitive. It is also beneficial to define a separate controlling algorithm to act as a *coordinator* between two or more coupled models.

The XCF defines an ExecutionCoordinator class to encapsulate a simulation model's execution strategy (see Fig. 5). The

analyze method outlines the controlling algorithm used to perform the simulation. Subclasses of ExecutionCoordinator are expected to redefine this method to tailor the algorithm to the specific characteristics of the model. Generally, this will involve creating Solver objects which are used to solve the model's equations. ExecutionCoordinator keeps a reference to a CompositeEngElement object, which represents the model being controlled by the ExecutionCoordinator. Thus, each assembly of EngElements represented by a CompositeEngElement can have its own ExecutionCoordinator and its own Solver. This allows nesting of Solvers and ExecutionCoordinators in a hierarchical model. A nested approach, such as this, is useful (and possibly) essential in multimodels. It allows models with different time scales to be handled by different solvers, which is important in dealing with mixed CFD and cycle models.

The XCF design defers the creation of solvers to the developer of the ExecutionCoordinator. However, to help developers identify EngElement models which declare equations to be used in a solver, the XCF defines the Solvable interface and a Solver class. Solvable is a Java interface which defines no method and no constants. It is a *marker* interface, which can be used to identify classes of objects (not only EngElements) which have equations to be solved. The design allows Onyx to identify objects which are to be handled by a solver, but lets developers decide how the data in those objects should be accessed. Class Solver is an abstract class which defines a single concrete method, buildSolvableList. This method returns a list of Solvable EngElements (i.e., instances of EngElement classes implementing the Solvable interface) for the CompositeEngElement it references (see Fig. 5).

6 Example

In Part 1, we developed a set of component-based models to simulate the E^3 gas turbine engine. We now utilize those component to graphically build and execute a distributed gas turbine simulation. The first step in the process is to add the ''Info class'' data—Customizer and EngElementInfo classes—to the component classes defined in Part 1. We describe the process only for the Compressor component; the other components are similarly modified.

6.1 Compressor Customizer. A Customizer for the CompressorModel was created to allow users to interactively change the Domain Model data when the simulation is run using Onyx's Visual Assembly Framework. The CompressorCustomizer class extends BasicTabbedCustomizer (see Fig. 3):

```
public class CompressorCustomizer
   extends BasicTabbedCustomizer {
```

When the CompressorCustomizer is instantiated by the Visual Assembly Framework, its constructor calls the buildTabbedPane method to create four CustomizerPage objects:

```
public CompressorCustomizer( ) {
    buildTabbedPane( );
  }
  private void buildTabbedPane( ) {
    CustomizerPage page1 = new CompressorGeneralPage( );
    addTab (page1, ''General'');
    CustomizerPage page2 = new CompressorDesignPtpage( );
    addTab (page2, ''Design point data'');
    CustomizerPage page3 = new CompressorInitOpPtPage( );
    addTab (page3, ''Initial operating point data'');
    CustomizerPage page4 = new CompressorTransientControlPage( );
    addTab (page4, ''Transient control'');
  }
}
```

Four new CustomizerPage subclasses were defined: CompressorGeneralPage, CompressorDesignPtPage, CompressorInitOpPtPage, and CompressorTransientControlPage. CompressorGeneralPage has an input fields for setting the name of the Compressor object. CompressorDesignPtPage has input fields for defining design point values for mass flow rate, adiabatic efficiency, stator geometry angle, and stator angle bias. The CompressorInitOpPtPage defines the initial operating point value of the stator angle. The CompressorTransientControlPage provides a TransientControlSchedule editor to define how the stator geometry angle varies during a transient. Using this CustomizerPage, the user creates *time-value* data points which are then added to the schedule. These values are displayed textually in a list on the page, and graphically by an XY plot of the value as a function of time. Linear interpolation is used to determine the stator angle value at times between the specified data points. The CompressorCustomizer is shown in the lower-left corner of Fig. 1.

6.2 CompressorInfo. As described in section 3.2, the Visual Assembly Framework composes SchematicIcons based on information defined by the EngElementInfo interface (see Fig. 2). The CompressorInfo class implements this interface:

```
public class CompressorInfo
   implements EngElementInfo {
```

The getEngElementDescriptior method returns an EngElementDescriptor object which encapsulates information about the Compressor. This information includes the Compressor class name, display name, and the class name of the CompressorCustomizer.

```
public EngElementDescriptor getEngElementDescriptor( ) {
  Class compClass = Class.forName (packageName+''Compressor'');
  Class customizerClass = Class.forName (packageName+''CompressorCustomizer'');
  EngElementDescriptor c = new EngElementDescriptor (compClass, customizerClass);
  c.setDisplayName (''Compressor'');
  return c;
}
```

CompressorInfo's `getPortDescriptor` method returns information concerning the number, type, display name and location of the VPorts to be included in the SchematicIcon. As described in Part 1, section 5.2, Compressor defines two OneDFluidPort objects and a single OneDMechPort object.

```
public PortDescriptor [ ] getPortDescriptors( ) {
  PortDescriptor fluidInPortDesc = new PortDescriptor(
      ''Fluid In,'' // Display name
      ''Fluid flow into Compressor,'' // short description
      PortDescriptor.IN_ONLY, // I/O type
      PortDescriptor.WEST); // Initial location

  PortDescriptor fluidOutPortDesc = new PortDescriptor(
      ''Fluid Out,'' ''Fluid flow out of Compressor,''
      PortDescriptor.OUT_ONLY, PortDescriptor.EAST);

  PortDescriptor mechInPortDesc = new PortDescriptor(
      ''Shaft In,'' ''Shaft into Compressor,''
      PortDescriptor.IN_ONLY, PortDescriptor.SOUTH);

  PortDescriptor[ ] ports = {fluidInPortDesc, fluidOutPortDesc, mechInPortDesc};
    return ports;
}
```

The `getPortDescriptor` method creates three PortDescriptor objects, each parameterized with appropriate Strings describing the, display name, short description, type, and location of the Ports. Each PortDescriptor object is added to the ports array and returned. The PortDescriptors are extracted by Onyx to build VPorts. Note that the order in which the PortDescriptors are added to the ports array is *extremely* important; the order must correspond to the order in which the Port objects were added to the ports Vector in the Compressor class. This is a consequence of the loose coupling between the UI objects and the EngElement objects. The only coupling between the Compressor class and the UI objects is the `getEngElementInfo` method defined by the EngElement interface (see Part 1, Fig. 3). In the Compressor class, this method appears as:

```
public EngElementInfo getEngElementInfo( ) {
   return (new CompressorInfo( ));
}
```

The other methods defined by EngElementInfo interface are also implemented in the CompressorInfo class, but are not listed here for brevity.

6.3 The Compressor JAR File. The Compressor, CompressorModel, DataTable, OneDFluidPort, OneDMechPort, CompressorCustomizer, CompressorGeneralPage, CompressorDesignPtPage, CompressorInitOpPtPage, Compressor TransientControlPage, CompressorInfo, and miscellaneous utility classes needed by Compressor (such as the TransientControllerSchedule class), were compiled using the Java compiler, *javac*, on a Silicon Graphics workstation (JDK 1.1.6). The resulting bytecode (.class) files were archived along with Icons representing the Compressor using the Java Archive (*jar*) tool following the description of section 3.4.

6.4 Setting Up the Example Simulation. Customizers and ''Info Classes'' for each of the EngElement classes (i.e., Bleed, Combustor, Duct, etc.) described in Part 1 were developed in a similar manner as for the Compressor. Each EngElement, its Domain Model, Customizer, and miscellaneous classes were archived into JAR files as just described. With the exception of the Combustor JAR file, all of the JAR files were copied to the machine running Onyx, and put in Onyx's *jar* directory. Jar files located in this directory are automatically loaded into Onyx when it is started.

The Combustor JAR file was placed on the University of Toledo Mechanical, Industrial and Manufacturing Engineering Department's Web server. The Combustor JAR file was then downloaded into Onyx and instantiated during model construction to demonstrate the ability to load component models from the Web. To further demonstrate Onyx's distributed capabilities, a Bleed EngElement was distributed to a remote machine and connected to the local machine running Onyx using RMI. The EngElement class, called RemoteBleed was placed on the remote machine, a DEC Alpha 255/300, located on the University of Toledo's College of Engineering network, and compiled using DEC's Java compiler. A simple server was written to export the RemoteBleed object to the RMI Registry. A RmiDistMgr class was written to handle the connection and referencing services using the RMI protocol. Because RMI treats remote objects as types of `java.rmi.Remote` rather than `java.lang.Object`, a wrapper called RmiProxyEngElement was written to wrap the remote EngElement proxy on the local machine. This class was defined as a subclass of DefaultEngElement, and thus could be loaded into Onyx. The class maintained a reference to the remote object proxy, and delegated any requests to the proxy.

6.5 Constructing the Engine Model. Using the local, remote and Web-based EngElements, a model representing the overall turbofan engine structure was constructed in the Visual Assembly Framework (see Fig. 1). Three sub-models were defined: Core, Turbofan, and Main.

The Core sub-model, represents the high-pressure core subsystem, and includes the High-Pressure Compressor, Combustor, High-Pressure Turbine and the High-Speed Shaft. Fluid connections were defined by instances of BasicFluidVConnector and mechanical connections were defined by instances of BasicMechanicalVConnector. Connections to the Turbofan sub-model were defined by promoting the input VPort on the High-Pressure Compressor and the exit VPort on the High-Pressure Turbine. Mix-

ingVolumes were placed between non-Shaft objects to define boundary conditions and handle inter-component fluid dynamics.

The Turbofan sub-model represents the low-pressure subsystem, and includes the Fan, Core, Low-Pressure Turbine, BypassDuct, Nozzle, and Low-Speed Shaft. Again, MixingVolumes were placed between non-Shaft objects to define boundary conditions and handle inter-component fluid dynamics. Three Bleed components were used to provide cooling air from the High-Pressure Compressor to the High-Pressure Turbine, Low-Pressure Turbine, and customer (over-board). The Bleed flow is extracted from MixingVolume 3, located downstream of the High-Pressure Compressor. Connections between the Bleed components and the other components were defined by promoting the output Port of each Bleed. These promoted Ports appear on the boundary of the Core SchematicIcon in the Turbofan schematic.

In the Main SchematicFrame, the complete engine model is represented by the Turbofan SchematicIcon which is connected at its inlet and exit to the Environment. This represents the placement of the engine in the environment, or in a test-cell.

6.6 Controlling the Simulation. An ExecutionCoordinator subclass, ExampleExecCoordinator (see Fig. 5), was created to define the controlling algorithm for executing the example simulation. For simplicity, only a single ExecutionCoordinator is used in the simulation, and it was (by default) associated with the top-level CompositeEngElement (which in this case is called Main). The coordinator defines steps for initializing the EngElements, performing a steady-state analysis at the engine's initial operating point, and a transient excursion from that point. The controlling algorithm is defined in the overridden `analyze` method:

```
void analyze(Object obj) {
    Solver aSteadyStateSolver = new NewtonRaphsonSolver( );
    Solver  aTransientSolver = new ExampleTransientSolver( );
    compositeEngElement.init( );
    aSteedyStateSolver.performSteadyStateAnalysis(compositeEngElement obj);
    aTransientSolver.performTransientAnalysis(compositeEngElement, obj);
}
```

For the example, two subclasses of Solver were created (see Fig. 5). The steady-state analysis was managed by an instance of NewtonRaphsonSolver, which defines the method `performSteadyStateAnalysis` to converge the system to steady-state using a Newton-Raphson numerical method. The transient analysis was managed by an instance of ExampleTransientSolver, which defines the method `performTransientAnalysis` to run the transient.

The `init` message is sent to the top-level CompositeEngElement, `compositeEngElement`, which in turn calls `init` on each of its children EngElements (see Part 1, Fig. 3). The `init` method is defined differently by each EngElement in the example. In most cases, it is used to determine *correction coefficients*, which are used for numerical stability in components which read data from performance maps [18].

6.7 Determining Steady-State Engine Balance. Before proceeding to the transient analysis, the simulation first attempts to drive the engine to balanced (steady-state) conditions at the initial operating point. This ensures that the engine model is in a consistent, physically valid operating state before beginning the transient.

The NewtonRaphsonSolver class uses a numerical algorithm based on the Newton-Raphson method [19], to drive the system to convergence. In order to apply this method, the NewtonRaphsonSolver class must identify the EngElements in the system that define differential equations, and obtain their derivative terms. EngElements defining differential equations, which are to be used in a Solver, are identified by implementing the Solvable interface (see Fig. 5). In the example engine system, the Duct, MixingVolume, and Shaft classes implement the Solvable interface. Each class implementing the Solvable interface defines `getEquation` and `setEquation` methods that are used to get/set instances of the Equation class. The Equation class, as its name implies, represent differential or algebraic equations. Instances of Equation are used to encapsulate the independent-dependent relationships of the domain model's equations. The system of equations is obtained by using the abstract Solver class' `buildSolvableList` method, which returns a list of Equations in the complete engine model. For the example engine system, this returns the following list of EngElements; `Duct BYPASS`, `MixingVolumeMV13`, `MixingVolumeMV3`, `MixingVolumeMV4`, `ShaftHSS`, `MixingVolumeMV41`, `Shaft LSS`, `Duct AFTERBURNER`, `MixingVolumeMV6`, and `MixingVolumeMV7`.

The NewtonRaphsonSolver object uses the `getDependent` method on each Solvable EngElement in the solvable list, and applies the Newton-Raphson algorithm to the system of equations to compute new values of the independent terms. These are placed back into their respective equations using the `setIndependent` method. The effects of the new independent terms on the system are evaluated by ''running'' the engine. The solver sends the run message to `compositeEngElement` and its `runSequence` method is called, which in turn calls run on each of its children EngElements (see Part 1, Fig. 3).

The execution order of the children EngElements is important in the example system. Calculations of certain fluid properties, such as fuel-air ratio, and enthalpy are dependent on the fluid conditions in the ''upstream'' component. Thus, the order is defined by the user for the entire model using the graphical sequence editor (see Fig. 1). For the example, the order is: `ENVIRONMENT` → `LPC` →`HPC`→`BYPASS`→`MV13`→`OV_BLEED` →`HI_BLEED`→`LO_BLEED`→`COMBUSTOR`→`MV3`→`HPT`→ `MV4` →`HPT_ROTOR`→`LPT`→`MV41`→`LPT_ROTOR`→`AFTERBURNER` →`MV6`→`NOZZLE`→`MV7`

6.8 Transient Engine Analysis. In unsteady simulations, the independent variables, $\mathbf{x}(t)$, are functions of time, and may be implicitly integrated to predict their value at a future time based on known past values. This is typically done using a family of numerical methods known as Predictor-Corrector methods [20]. In these methods, an initial estimate of the independent variable at the end of a discrete time step, Δt, is estimated using a Predictor equation. The system is evaluated using the estimated independent to calculate the dependent derivative term. The derivative is then used in a Corrector equation that computes a new estimate of the independent at the end of the time step. The predictor estimate and the corrector estimate are compared. If the difference is less than

some user-defined tolerance, then the equation is converged. If not, the Corrector is applied again using the most recent estimate of the independent.

The same approach as described for the steady-state case was used for the transient, except in this case, the time value was varied and the independent calculated at the new time value. An algorithm employing the Euler Implicit Predictor-Corrector method was used to estimate new values for each independent in the set of equations.

6.9 Execution Coordinator Customizer. An ExecutionCoordinatorCustomizer was written to provide interactive control over the simulation, and to allow users to change the Solver input data. The approach is essential the same as was described for the Compressor Customizer class. Three subclasses of CustomizerPage were created: ExampleControlPage, NewtonRaphsonSolverPage, and ImplicitEulerSolverPage. The ExecutionCoordinatorCustomizer is seen at the top center of Fig. 1. The ExampleControlPage has an input field for users to set the transient duration, and buttons to start and halt the simulation. Two progress bars are provided to give the user visual feedback about the progression of both the steady-state balancing and the transient simulation. The NewtonRaphsonSolverPage and ImplicitEulerSolverPage pages provide input fields for setting specific control values for the operation of each solver.

6.10 Running the Simulation. After the engine model was constructed, each EngElement's design and initial operating point data were fully defined, the Solver parameters were set, and the simulation was started by pressing the ''Run Simulation'' button in the ExecutionCoordinatorCustomizer's *Control* page. The engine was balanced successfully at the initial operating point in three iterations using the Newton-Raphson solver, and a two second transient was then run. During the transient, fuel flow rate in the Combustor was increased linearly from 0.37 to 1.7 lbm/s. (see bottom center of Fig. 1) During that time, the Fan and HP Compressor stator angles were adjusted according to the TransientControlSchedule defined in the CompressorCustomizer to improve flow into the blades. Component parameters for each component in the simulation model can be plotted using Onyx's plotting facilities. For example, Fig. 1 (bottom right) shows a plot of the stagnation pressure (psi) for each of the Mixing Volumes in the model.

7 Conclusion

The Onyx domain framework described in this paper provides an ensemble of framework components which, together, form an integrated framework for propulsion system simulation. Onyx utilizes advanced object-oriented technologies such as object-oriented frameworks and design patterns to produce a reusable component-based architecture which can be extended and customized to meet future application requirements.

Onyx promotes the construction of aerospace propulsion systems, such as jet gas turbine engines, in the following ways. First, it provides a Common Engine Model which (i) encapsulates the hierarchical nature of the physical engine model, (ii) is capable of housing multidisciplinary and multifidelity analysis models, and (iii) enforces component interoperability through a consistent interface between components. Second, it enables the construction of engine models and customization of the simulation at a high level of abstraction through the use of visual representation in the Visual Assembly Framework. Third, it provides a general, yet powerful, abstraction mechanism for distributing the computational requirements of a gas turbine simulation using services such as RMI, CORBA, Voyager or DCOM. And finally, it supports the integration of user-defined execution strategies based on the solvers being used, as well as the objective of the simulation.

Acknowledgments

The work described in this paper was made possible by funding from the NASA Glenn Research Center Computing and Interdisciplinary Systems Office, and the University of Toledo. We would like to thank Greg Follen at NASA Glenn for his continued support.

References

[1] Prieto-Díaz, R., 1990, *Domain Analysis: An Introduction*, ACM SIGSoft Software Engineering Notes, **15**, No. 2, pp. 47–54.
[2] Batory, D., McAllester, D., Coglianese, L., and Tracz, W., 1995, ''Domain Modeling in Engineering of Computer-Based Systems,'' *Proc. of the 1995 International Symposium and Workshop on Systems Engineering of Computer Based Systems*, Tucson, Arizona.
[3] Gamma, E., Helm, R., Johnson, R., and Vlissides, J., 1995, *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison/Wesley, Reading, MA.
[4] Schmidt, D. C., 1997, ''Applying Design Patterns and Frameworks to Develop Object-Oriented Communications Software,'' *Handbook of Programming Languages*, Vol. I, Salus, I. P., ed., MacMillian Computer Publishing.
[5] Schmidt, D. C., 1995, ''Experience Using Design Patterns to Develop Reusable Object-Oriented Communication Software,'' Commun. ACM, **38**, No. 10, pp. 65–74.
[6] Johnson, R. E., and Foote, B., 1988, ''Designing Reusable Classes,'' J. Object-Oriented Prog., **1**, No. 2, pp. 22–35.
[7] Englander, R., 1997, *Developing Java Beans*, O'Reilly & Associates, Sebastopol, CA.
[8] Arnold, K., and Gosling, J., 1996, *The Java Programming Language*, Addison/Wesley, Reading, MA.
[9] Weiner, S. R., and Asbury, S., 1998, *Programming with JFC*, Wiley Computing, New York, NY.
[10] Curlett, B. P., Haas, A. R., and Naylor, B. A., 1995, ''Adaptive Graphical User Interface Framework for Object-Oriented System Simulations,'' NASA TM-106890.
[11] Holst, T. L., Salas, M. D., and Claus, R. W., 1992, ''The NASA Computational Aerosciences Program—Toward Teraflops Computing,'' AIAA Paper No. 92-0558.
[12] Wollrath, A., Riggs, R., and Waldo, J., 1996, ''A Distributed Object Model for the Java™ System,'' *The Second Conference on Object-Oriented Technology and Systems (COOTS) Proceedings*, pp. 219–231.
[13] *Voyager*, 1997, ''Voyager: The Agent ORB for Java,'' http://www.objectspace.com.
[14] Vinoski, S., 1997, ''CORBA: Integrating Diverse Applications Within Distributed Heterogeneous Environments,'' IEEE Communications, **35**, No. 2, pp. 46–55.
[15] Brown, N., and Kindel, C., 1998, ''DCOM: ''Distributed Component Object Model Protocol,'' Microsoft Corporation.
[16] Berners-Lee, T., 1996, ''WWW: Past, Present, and Future,'' Computer, **29**, No. 10, p. 69–77.
[17] Page, E. H., Buss, A., Fishwick, P. A., Healy, K. J., Nance, R. E., and Paul, R. J., 1998, ''The Modeling Methodological Impacts of Web-Based Simulation,'' *Proceedings of the 1998 SCS International Conference on Web-Based Modeling and Simulation*, pp. 123–128.
[18] Reed, J. A., 1993, ''Development of an Interactive Graphical Aircraft Propulsion System Simulator,'' MS thesis, The University of Toledo, Toledo, OH.
[19] Gerald, C. F., and Wheatley, P. O., 1984, *Applied Numerical Analysis*, Addison-Wesley, Reading, MA.
[20] Carnahan, B., Luther, H. A., and Wilkes, J. O., 1990, *Applied Numerical Methods*, R. E. Krieger, Malabar, FL.

Ufuk Senturk

Rogerio S. Lima

Center for Thermal Spray Research,
Department of Materials Science
and Engineering,
State University of New York at Stony Brook,
Stony Brook, NY 11794-2275

Carlos R. C. Lima

UNIMEP Methodist University of Piracicaba,
Technology Center,
Rod. Santa Barbara-Iracemapolis, Km 1,
Santa Barbara d'Oeste, S. Paulo, 13450-000,
Brazil

Christopher C. Berndt

Center for Thermal Spray Research,
Department of Materials Science
and Engineering,
State University of New York at Stony Brook,
Stony Brook, NY 11794-2275

Deformation of Plasma Sprayed Thermal Barrier Coatings

The deformation behavior of thermally sprayed partially stabilized zirconia (PSZ) coatings are investigated using Hertzian indentation and four-point bend testing, with in situ acoustic emission monitoring. The experimental deformation curves, together with the corresponding acoustic emission responses and the fracture properties of the material are used in defining the deformation characteristics of the coating (ceramic overlay with metallic bond coat where applicable) and substrate composite system. Experiments are aimed in examining the influence of the bond coat and the coating properties on the form of deformation. Substrate temperature and pauses during spraying are demonstrated to strongly effect the coating properties and the resulting fracture/failure characteristics of the composite system. [S0742-4795(00)02503-5]

Keywords: Plasma Spray, PSZ, Coating, Acoustic Emission, Mechanical Property

Introduction

Thermal Barrier Coatings (TBCs) have been widely used in gas turbine and diesel engines [1]. The most common material used as an insulator in such systems has been partially stabilized zirconia (PSZ), mainly for its low thermal conductivity and diffusivity combined with its high temperature stability. The general practice in preparing thermal spray TBCs is to spray the ceramic overlay onto a metallic bond coat layer, generally by plasma spraying for large components. Plasma sprayed TBCs experience high thermal stresses under service conditions due to cyclic thermal changes. Thermal stresses are generated mainly due to the thermal expansion coefficient mismatch between the ceramic coating and the metallic substrate. From the viewpoint of structural integrity and mechanical behavior, the presence of a metallic bond coat layer between the ceramic and substrate can reduce the effect of such stresses [2]. In addition, plasma sprayed TBCs are subjected to residual stresses that arise from the rapid cooling of molten or partially molten droplets impacting on the cool substrate, from thermal gradients, and from solid state transformations of the coating material. Hence, controlling the substrate temperature during deposition has a significant influence on the nature and magnitude of residual stresses [3–5].

Several approaches have been used to improve the mechanical properties, and thermal cycle and oxidation resistance behavior of TBCs [6–8]. These include (i) optimizing Young's modulus, (ii) reducing the thermal expansion coefficient mismatch between the coating and substrate, and (iii) increasing the oxidation resistance, especially near the substrate interface, at high temperatures. To better understand the coating formation and failure mechanisms, the influence of spray conditions and TBC system design have been characterized by a number of distinct methods. Acoustic Emission (AE) technology has been widely used as a non-destructive method in characterizing the cracking and failure behavior of the coatings. The technique also has potential for the quality control and in-service monitoring of coatings during application or production. The method has mostly been used to understand the failure mechanisms during thermal cycling [9,10]. Less effort has been spent in characterizing the deformation and failure behavior at room temperature conditions, under mechanical stresses. Among the few studies, experiments have monitored the *in situ* cracking of samples during three-point [11,12] and four-point bend tests [13–15], tensile adhesion tests [16–20], and indentation [21] tests.

In this paper, the mechanical properties and cracking features of yttria stabilized zirconia (YSZ) ceramic coatings, either with or without a bond coat (NiCrAl) layer, plasma sprayed on plain carbon steel substrates are investigated. Four point bending and Hertzian indentation tests have been used to evaluate the mechanical response, while acoustic emission (AE) has been used to *in situ* monitor the cracking behavior during these tests. The bend tests are used to provide more quantitative results with respect to the mechanical properties of the coating-substrate system. The indentation tests are used for the simple determination of indentation loads that lead to failure and to gather general information about the deformation and adhesion behavior, which complement the information obtained from the bend tests. The mechanical tests are performed on coatings that were produced under processing conditions identified as either ''cooling'' or ''no cooling'' during the spray procedure are examined. Moreover, two types of coating structures are distinguished with coatings produced by a continuous spraying procedure and by a paused scheme.

Experimental Procedures

Materials and Spray Conditions. The coatings were air plasma sprayed (APS) using the conditions given in Table 1 with the plasma spray gun (Metco 3MB, Westbury, NY) mounted on a six-axis articulated robot (GMF Fanuc, Model S400). The feedstock material was yttria stabilized zirconia (YSZ), of average particle size of 56.2 μm for the top coat and NiCrAl for the bond coat, both externally injected into the plasma at a rate of 40 g/min using nitrogen as the carrier gas. Spraying was performed on mild

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-348. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

Table 1 Plasma spray conditions

	YSZ	NiCrAl
Gun Type	Metco 3MB	Metco 3MB
Current	600A	500A
Voltage	70V	70V
Primary Gas	Ar (40 *l*/min)	Ar (40 *l*/min)
Secondary Gas	H_2 (8 *l*/min)	H_2 (11 *l*/min)
Powder Carrier Gas	N_2 (3.5 *l*/min)	N_2 (3.65 *l*/min)
Spray Distance	100mm	100mm

steel substrates of 60×7×2.54 mm size, providing it with the bond coat layer, where applicable, and the YSZ topcoat. Prior to spraying, the surfaces to be sprayed were grit blasted and washed in ethyl alcohol. For spraying, the samples were placed on a carousel with six substrates sprayed at a time. The rotation speed of the carousel was set at 160 rpm, while the robot was programmed to provide a vertical traverse distance of 90 mm at 12 mm/s. For samples where a bond coat layer was produced, the robot was set to repeat 10 uninterrupted passes, giving a thickness of 150–200 μm. For the YSZ topcoat, the spray gun was set to repeat a total of 100 passes, giving a thickness in the 900–1100 μm range. During the application of the ceramic coating, samples were coated in a continuous as well as in a paused schedule. For the continuous coatings, the spraying was performed without any interruptions until the desired thickness was reached. For the paused coatings, the spraying was stopped for 2 times (both after 40 passes) where the samples were allowed to cool to ~100 °C; afterward the coating process was resumed until the desired thickness was achieved. It should be noted that for both the continuous and paused coatings, the thicknesses of the bond coat and ceramic layer was the same.

Coatings that were prepared continuously or with pauses were sprayed either with or without external cooling applied using pressurized air jets to maintain a constant substrate temperature. The temperature of the substrate was determined using temperature indicating liquids in different measuring ranges (Omegalaq, Omega Engineering, Stamford, CT) and was confirmed using a hand held infrared temperature detector. The temperatures measured were in the 400–500 °C range for the ''cooled'' samples and was above 500 °C for the ''non-cooled''. It is also important to note the difference in the heating rates when making a distinction between the two cases. Samples having ''bond coat'' and ''no bond'' coat were made for the different cases described above.

Mechanical Testing and Acoustic Emission

Four Point Bending Tests. The samples were subjected to four-point bending (20 mm inner span and 40 mm outer span) using a servo-hydraulic testing machine (Model 8502, Instron, Canton, MA), with the crosshead speed set to 10 μm/sec and the load and displacement recorded for each measurement. A maximum load of 1 kN was applied to all the samples. A compliance calibration for the experimental arrangement was made using the Instron LVDT and an external LVDT (Model CD375-500, Macro Sensors, Pennsauken, NJ) that was placed close to the pin supports. This enabled a better measurement of displacement by calibrating the crosshead movement. During the tests, a piezoelectric AE transducer (pico-transducer, Physical Acoustics, Princeton, NJ) with a resonance frequency of 250 kHz was placed on the coating. The sensor was attached to a preamplifier (model 140B, Hartford Steam Boiler Inspection Technologies, Sacramento, CA) and a preamplifier filter (model FL12Y, Hartford Steam Boiler Inspection Technologies, Sacramento, CA), with an amplification of 40 dB and the frequency range analyzed at 10 kHz-1 MHz. The signal output was processed using an AET 5500 system (Hartford Steam Boiler Inspection Technologies, Hartford, CT), which was connected to a redundant PC to record the acoustic emission (AE) responses. Procedures with regards to the experimental setup and AE signal processing are available elsewhere [14]. A total of two samples were tested for each condition. Prior to testing, the sample surfaces were polished to a 1 μm diamond finish and the specimens were stored in a desiccator. Blank steel substrates were also tested to examine the influence of the coating on the acoustic emission activities so that the acoustic emission contribution from the substrate could be taken into account.

Hertzian Indentation Tests. Hertzian (spherical) indentation tests were performed using an experimental set-up built on a servo-hydraulic testing machine (Model 8502, Instron, Canton, MA). This enabled load and displacement data to be collected during the indentation cycle, with the load detected using a load cell and the displacement using an external LVDT (Model CD375-500, Macro Sensors, Pennsauken, NJ). The LVDT in this case was placed close to the indenter tip so that the displacements from the indenter penetration can be more accurately measured. The indentations were performed using a 3.175 mm (1/8″) diameter WC ball. The crosshead speed during the indentation cycle was set to 10 μm/s and a maximum load of 3 kN was applied to all the samples (Note: some coatings failed before this maximum load). The same AE transducer and data analysis system was used for both the four-point bending and indentation tests.

Results and Discussions

Four-Point Bending and Acoustic Emission. The stress-displacement curves for the ''cooled'' and the ''non-cooled'' coated samples are shown in Fig. 1(*a*) and Fig. 1(*b*), respectively. Stress was calculated using the maximum stress developed within the inner span, according to the following equation:

$$\sigma = \frac{3PD}{h^2 d},$$

where P is the applied load, D is the distance between the loading and support bar (10 mm), h is the thickness, and d is the width of the bar. Displacement was determined using the values measured from the crosshead motion (excluding the set-up compliance). The calculated stress is the tensile force that is present on the ceramic surface (also referred to as the outer fiber stress). It should be noted that the calculation of the stress assumes the coating-substrate composite system as one material, allowing the applied bending loads to be normalized with respect to the variations in the sample thickness and width. Hence, a direct comparison of these cooled and non-cooled samples is only possible from the load-displacement curves; but in such a case different sample dimensions will need to be taken into account.

All the curves in Fig. 1 show an initial linear range, which then changed to a non-linear portion, indicative of an elastic-plastic response. The initial non-linear ''toe'' region can also be recognized, and is attributed to the arrangement of the loading and support bars on the sample. This behavior of the composite system arises since the composite consists originally of a steel substrate and in some instances a metallic bond coat. Hence, the transition from the elastic to plastic regions can be related to the behavior of the metal substrate, with the deformation characteristics of the metal substrate dominating after yielding.

The presence of the ceramic coating, together with the bond coat layer where applicable, has a significant influence on the stress-displacement curve prior to yield. This can be recognized by the presence of a peak transition region between the elastic and plastic portions when the coating is present on the substrate, suggesting that the coating is structurally contributing to the material responses of the composite structure until yielding. Therefore, this behavior indicates that the coating is inhibiting yield of the composite system with respect to the usual behavior of the metal substrate where there is a smooth transition region. Thus, the coating appears to reinforce or strengthen the system and the yield point is elevated. Accordingly, the peak stress can be related to the failure

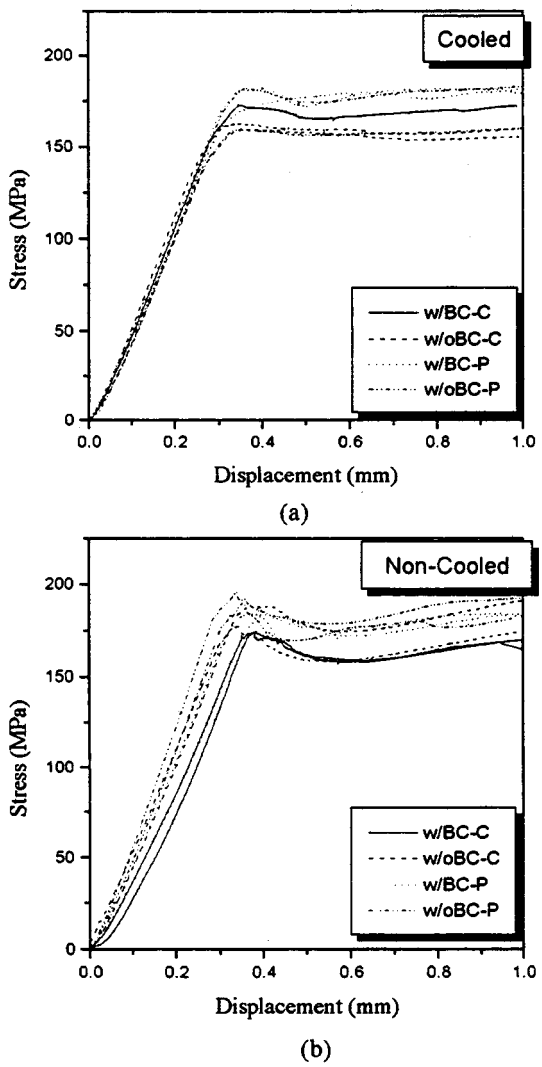


Fig. 1 Stress-displacement curves from four-point bending tests (*a*) for cooled and (*b*) non-cooled samples. The figures include samples having bond coat ("w/BC") or no bond ("w/oBC") coat layer which are either sprayed continuously ("C") or with pauses ("P").

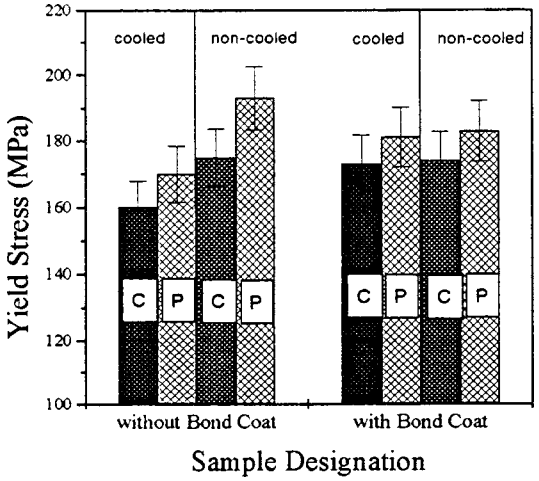


Fig. 2 Yield stress of the coating system determined from the four-point bending curves. The results are given for samples having a bond coat and no bond coat layer, which are sprayed either continuously ("C") or with pauses ("P").

of the coating [15]. It should also be noted that the peak load behavior is more pronounced for the ''non-cooled'' coatings.

The changes in the yield stress for samples sprayed with and without a bond coat layer, for the cooled and noncooled cases having a continuous and paused spray structure is shown in Fig. 2. A three-level analysis of variances (ANOVA) test on the data indicated that the cooling and pausing features are statistically significant factors that have an effect on the yield stress results. The influence of these variables will also be shown and discussed for the AE results. A significant feature which is not highlighted by the statistical analysis results but can be recognized from Fig. 2 is that a larger variation in the average yield stresses is present for samples with no bond coat layer (can be observed as an increasing trend in the figure) while the bond-coated samples have more equivalent results. This indicates that (*i*) the presence of the bond coat has a critical influence on the results, and (*ii*) the application of cooling or no cooling and any pauses during spraying has a more significant effect on samples having no bond coat layer applied. This behavior can be explained on the basis of the role of the bond coat layer, where, among its other features, it is also known [22] to act as a compliant layer in decreasing the thermal expansion mismatch between the ceramic coating and the metal substrate. The benefit from this behavior, which is normally intended to be dominant at high temperatures, is that the ceramic overlay is isolated from high strain and hence, is prevented from spallation [22]. For repeated thermal cycling, this continues until the bond coat can not plastically deform, i.e., creep. This reason for this behavior has been attributed to the increase in the strain energy release rate for buckling spallation and a corresponding decrease in critical flaw size for failure due to the formation of an oxide layer [23,24]. It is also argued that the depletion of aluminum from the bond coat results in an oxide layer other than α-alumina that have inferior mechanical properties and lead to failure [25].

Stresses arising in the ceramic layer during the initial cooling, on the other hand, are reported to be compressive in the ceramic [22] near to the ceramic-bond coat interface. This behavior depends on the other processing parameters; specifically on the substrate temperature [26,27]. The application of cooling or no cooling during spraying, thus, is related directly to the substrate temperature. Accordingly, cooling, whether it is applied or not, can be expected to influence the ceramic coating and the interfaces that it abuts (either the bond coat or substrate). It will be shown with the indentation analysis that the influence of cooling has a greater influence on the bond coat layer and the bond coat-ceramic interface.

Another significant feature that is recognized from Fig. 2 is that the samples sprayed with pauses all tend to exhibit a higher yield stress compared to the continuously sprayed coatings of the same thickness. It is quite likely that there are stresses concentrated at the interfaces that are formed during the spray interruption. This is illustrated in Fig. 3 where the crack which has originated from the surface has deflected at the interface where the coating cycle was paused that is close to the bond coat layer. A similar behavior of thermal spray coatings under indentation stresses was shown in another study [28,29]. Apparently, the presence of these regions contributes to the improvement in the low temperature integrity of the composite system.

The total AE events recorded during the linear elastic portion of bending for the various samples are shown in Fig. 4. The elastic region AE events are investigated to study the deformation characteristics of the ceramic coating since after yielding the ceramic coating is known to fail during bending. For the continuous sprayed coatings, cooling and the presence of a bond-coat layer causes lower AE activity during bending compared to the non-cooled coatings. Similar behavior was shown and discussed in another study for a different coating thickness [15], where residual

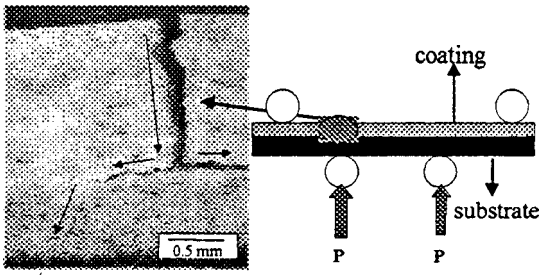


Fig. 3 Side view of sample deformed using four-point bending, showing cracks that have originated from the surface at the high stress region and have propagated until it was deflected at the interface formed by the pauses during spraying

stresses and absorption of AE activity [10] were factors responsible for this behavior. When the coating is sprayed with pauses, the most significant feature recognized is the decrease in the total AE activity for the non-bond coated samples and an increase for the bond-coated samples, both compared with the continuously sprayed coatings. Also, the cracking events for the non-continuously sprayed coatings with and without bond coat are reasonably equivalent when compared with the continuous coatings. Hence, it can be concluded that discontinuities during spraying tend to create a layered structural formation (see Fig. 3), which causes cracking during elastic deformation in a similar way for both the bond coated and non-bond coated samples. At this point, it is difficult to explain this behavior using only the mechanical and acoustic responses of the samples. However, the layered formation in the coating, which is seen to strongly influence the mechanical characteristics of the coating, most likely relates to the stress distribution in the ceramic coating that is largely effected by the temperature variations caused by the pauses.

Hertzian Indentation and Acoustic Emission. The acoustic emission energy response for the cooled and non-cooled samples, with and without the presence of a bond coat layer, for coatings that are continuously or non-continuously sprayed is shown in Fig. 5. It should be noted that, although a total of six indentations were performed and analyzed for each condition, only one representative result was selected and is shown in the figure to illustrate the trends detected. One significant feature that is recognized from the curves is that there is a critical load (L_{cr}) after which the AE energy increases dramatically. This load has been correlated with the failure of the coating under the large stresses. Table 2 shows these loads for all the spraying conditions applied in this study.

The failure, where applicable, has been detected to be in the form of radial cracks emanating from the indentation area or as spallation/removal of the coating from the substrate. Note that not all the samples show this behavior. Coatings with a bond coat layer that were cooled during spraying did not fail until the maximum load (3 kN) applied in this study. This correlates well with the bending results discussed in the previous section as well as with previous studies [15]. It was noted in these discussions that the presence of the bond coat layer suppresses AE activity [10] due to surface compressive stresses inhibiting cracks. The indentation results, in support of the bend test results and application experience, verify that the bond coat layer improves the overall integrity of the TBC system.

On examining Fig. 5, it is observed that the critical load is higher when a bond coat layer is present for both the continuously and non-continuously sprayed coatings. This, being in good agreement with the bending results supports the well-accepted concept

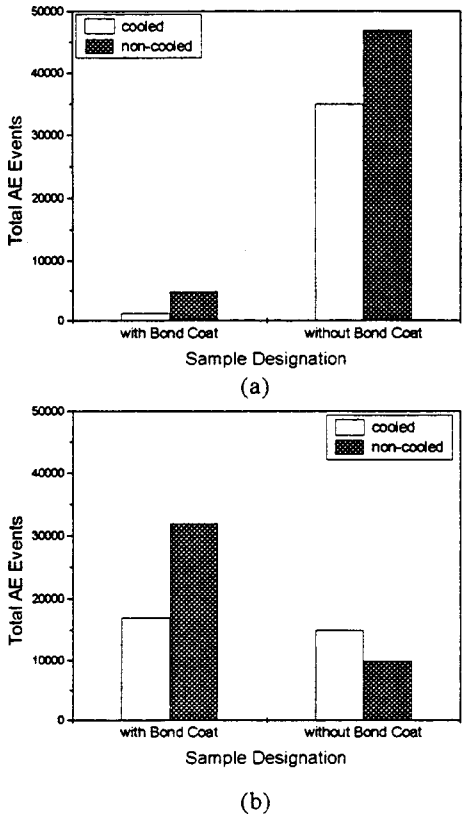


Fig. 4 Total number of AE events for four-point bending samples detected before yielding, for samples that are sprayed (*a*) continuously and (*b*) with pauses

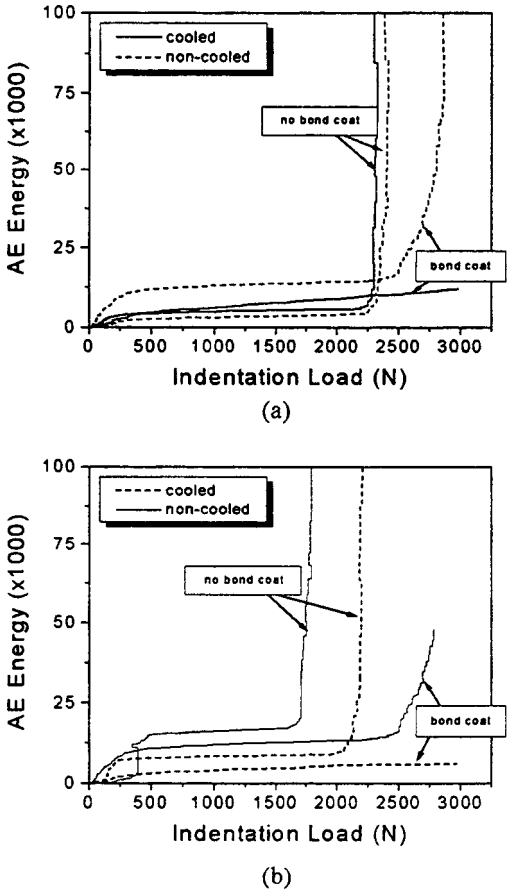


Fig. 5 AE energy versus indentation load for samples that are sprayed (*a*) continuously and (*b*) with pauses

Table 2 Critical load, "L_{cr}", defining the onset of larger AE activity

	Continuous sprayed	Paused sprayed
Cooled		
With bond coat	>3000	>3000
Without bond coat	2000	2200
Non-Cooled		
With bond coat	2500	2500
Without bond coat	1650	2250

that the bond coat is a major constituent in contributing to the integrity of the TBC system. It is known that during Hertzian indentation, the stresses below the indenter are compressive during loading while the surrounding area is in tension [30]. Hence, the coating below the indenter is compressed while its surroundings are pulled apart. Accordingly, when the loads are large enough, the coating fails mainly due to the tensile stresses that cause the pre-existing cracks to grow and the pores to coalesce. Thus, when the bond coat layer is present, such cracks are prevented from forming and growing, and confer good coating integrity.

The effect of discontinuities during spraying has been more significant for samples that have no bond coat, where decreases in the critical load, L_{cr}, can be noted; and has been especially significant for the non-cooled sample. Another characteristic for these samples is that both exhibit an initial step in the energy versus load curve at loads lower than 500N. The fact that the discontinuities during spraying cause a layered coating has been suggested in the previous discussions. This idea is once more strengthened with the indentation AE data, where more cracking is noted for these layered coatings. The initial step in the AE response can be correlated with the layered structure, most probably to cracks forming and growing between the layers. Also, cooling during spraying was shown to form a stronger coating in a former study [10], most probably due to surface compressive stresses. This is again supported with the larger L_{cr} for the cooled sample with the layered coating, compared with the non-cooled sample, both having no bond coat layer.

Conclusions

The yield stress from the bending tests and the indentation loads leading to coating failure together with the deformation characteristics of coating-substrate system are shown to be significantly influenced by external cooling, the spraying procedure (continuous or paused), and by the presence of a bond coat layer. It is shown that the bending deformation behavior of coatings can be used as a simple guide to gain information about the integrity of the coating-substrate composite systems and, more importantly, the characteristics of the coating deformation behavior. Using this analysis, changes in the mechanical response of the composite are observed with respect to influences of various spray parameters and processing methods. The combination of these parameters are factors that determine the mechanical properties and, hence, the failure characteristics of the coatings.

On the other hand, it is shown that the influence of processing parameters are more striking on the AE behavior. The AE analysis during bending and Hertzian indentation tests shows that the cracking phenomena in these samples change, depending on the processing parameters used. For bending studies, the deformation during the elastic region is significant in determining the fracture characteristics. For indentation, the critical load where the structure fails is determined and is shown to be related significantly to the spraying parameters without extracting information from the load-displacement behavior [28,29]. The overall analysis of the results indicate that coating adhesion and integrity and residual stresses play a major role on the influence of mechanical and deformation characteristics of the system as a function of any cooling application and presence of bond coat layer. The influence of the bond coat layer in improving the mechanical adhesion and coating integrity, the application of cooling to strengthen the coating, and obtaining a less flawed coating with a continuous spraying procedure are shown as significant characteristics of the coating that can be ascertained from bending and indentation tests.

Acknowledgments

C. R. C. Lima would like to thank FAPESP-Brazil for the financial support during his stay at SUNY Stony Brook. R. S. Lima acknowledges the support of an ITSA scholarship. This work was sponsored under NSF-MRSEC DMR, grant number 9632570.

References

[1] Pawlowski, L., 1995, *The Science and Engineering of Thermal Spray Coatings*, Wiley, New York.
[2] Taylor, R., Brandon, J. R., and Morrel, P., 1992, ''Microstructure, Composition and Property Relationships of Plasma-Sprayed Thermal Barrier Coatings,'' Surf. Coat. Technol., **50**, pp. 141–149.
[3] Clyne, T. W., and Gill, S. C., 1996, ''Residual Stresses in Thermal Spray Coatings and Their Effect on Interfacial Adhesion: a Review of Recent Work,'' J. Thermal Spray Technol., **5**, pp. 401–418.
[4] Clyne, T. W., and Gill, S. C., 1994, ''Investigation of Residual Stress Generation During Thermal Spraying by Continuous Curvature Measurement,'' Thin Solid Films, **250**, pp. 172–180.
[5] Kuroda, S., Dendo, T., and Kitahara, S., 1995, ''Quenching Stress in Plasma Sprayed Coatings and its Correlation with the Deposit Microstructure,'' J. Thermal Spray Technol., **4**, pp. 75–84.
[6] Zhou, Z., Eguch, N., and Ohmori, A., 1997, ''Microstructure Control of Zirconia Thermal Barrier Coatings by Using YAG Laser Combined Plasma Spraying Technique,'' *Thermal Spray: A United Forum for Scientific and Technological Advances*, C. C. Berndt, ed., ASM International, Materials Park, OH, pp. 315–321.
[7] Lima, C. R. C., and Trevisan, R. E., 1997, ''Graded Plasma Spraying of Premixed Metal Ceramic Powders on Metallic Substrates,'' J. Thermal Spray Technol., **6**, pp. 187–192.
[8] Tsui, Y. C., and Clyne, T. W., 1996, ''Adhesion of Thermal Barrier Coating Systems and Incorporation of an Oxidation Barrier Layer,'' *Thermal Spray: Practical Solutions for Engineering Problems*, C. C. Berndt, ed., ASM International, Materials Park, OH, pp. 275–284.
[9] Bordeaux, F., Moreau, C., and Saint Jacques, R. G., 1992, ''Acoustic Emission Study of Failure Mechanisms in TiC Thermal Barrier Coatings,'' Surf. Coat. Technol., **54/55**, pp. 70–76.
[10] Shankar, N. R., Berndt, C. C., Herman, H., and Rangaswamy, S., 1983, ''Acoustic Emission from Thermally-Cycled Plasma-Sprayed Oxides,'' Am. Ceram. Soc. Bull., **62**, pp. 614–619.
[11] Zhang, H. T., Zhou, X. K., Guan, K., Liao, B., and Cao, S., 1987, ''Acoustic Emission Research on the Fracture Behavior of Plasma-Sprayed Ni-Al Coatings During Bend Testing,'' Surf. Coat. Technol., **30**, pp. 115–123.
[12] Berndt, C. C., Robins, D., Zatorski, R., Herman, H., 1983, ''Fire Barrier Coatings for Protection of Aluminum Structures,'' *Proceedings of the 10th International Thermal Spraying Conference*, DVS Berichte, Dusseldorf, pp. 182–186.
[13] Lin, C. K., Berndt, C. C., Leigh, S., and Murakami, K., 1997, ''Acoustic Emission Studies of Alumina-13 percent Titania Free-Standing Forms During Four-Point Bend Tests,'' J. Am. Ceram. Soc., **80**, pp. 2382–2394.
[14] Lin, C. K., 1995, ''Statistical Approaches to Study Variations in Thermal Spray Coatings,'' Ph.D. thesis, State University of New York at Stony Brook, Stony Brook, NY.
[15] Senturk, U., Lin, C. K., Lima, R. S., Lima, C. R. C., and Berndt, C. C., 1999, ''Processing and Mechanical Properties of Plasma Sprayed Thermal Barrier Coatings,'' to be published in the 1999 United Thermal Spray Conference Proceedings.
[16] Berndt, C. C., Phucharoen, W., and Chang, G. C., 1984. ''The Mechanical Properties of Thermal Barrier Coatings Used for Gas Turbine Blades,'' Turbine Engine Hot Section Technology, NASA Workshop, pp. 155–166.
[17] Berndt, C. C., 1988, ''Cracking Processes in Thermally Sprayed Ceramic Coatings,'' Mater. Sci. Forum, **34-36**, pp. 457–461.
[18] Shankar, N. R., Berndt, C. C., and Herman, H., 1983, ''Structural Integrity of Thermal Barrier Coatings by Acoustic Emission Studies,'' *Proceedings of the 10th International Thermal Spraying Conference*, DVS Berichte, Dusseldorf, pp. 41–45.
[19] Shankar, N. R., Berndt, C. C., and Herman, H., 1983, ''Characterization of the Mechanical Properties of Plasma-Sprayed Coatings,'' *Materials Science Research Advances in Materials Characterization*, Plenum Press, New York, NY, **15**, pp. 473–489.
[20] Shankar, N. R., Berndt, C. C., and Herman, H., 1982, ''Failure and Acoustic Emission response of Plasma-Sprayed ZrO2-8wt percent Y2O3 Coatings,'' Ceram. Eng. Sci. Proc., **3**, pp. 772–792.

[21] Safai, S., Herman, H., and Ono, K., 1980, ''Acoustic Emission Study of Thermally Cycled Plasma Sprayed Oxide Coatings,'' *Proceedings of the 9th International Thermal Spraying Conference*, Netherlands Instituut voor Lastechniek, pp. 129/132.
[22] Bennet, A., 1986, ''Properties of Thermal Barrier Coatings,'' Mater. Sci. Technol., **2**, pp. 257–261.
[23] He, M. Y., Evans, A. G., and Hutchinson, J. W., 1996, ''Interface Cracking Phenomena in Constrained Metal Layers,'' Acta Mater., **44**, No. 7, pp. 2963–2971.
[24] He, M. Y., Evans, A. G., and Hutchinson, J. W., 1998, ''Effects of Morphology on the Decohesion of Compressed Thin Films,'' Mater. Sci. Eng., A, **A245**, No. 2, pp. 168–181.
[25] Brindley, W. L., and Whittenberger, J. D., 1993, ''Stress Relaxation of Low Pressure Plasma-Sprayed NiCrAlY Alloys,'' Mater. Sci. Eng., A, **A163**, No. 1, pp. 33–41.
[26] Bianchi, L., Luchesse, P., Denoirjean, A., and Fauchais, P., 1995, ''Zirconia Splat Formation and Resulting Coating Properties,'' *Proceedings of the 8th National Thermal Spray Conference*, C. C. Berndt and S. Sampath, eds., ASM International, Material Park, OH, pp. 261–266.
[27] Bianchi, L., Luchesse, P., Denoirjean, A., Fauchais, P., and Kuroda, S., 1995, ''Evolution of Quenching Stress During Ceramic Thermal Spraying with Respect to Plasma Parameters,'' *Proceedings of the 8th National Thermal Spray Conference*, C. C. Berndt and S. Sampath, eds., ASM International, Material Park, OH, pp. 267–271.
[28] Pajares, A., Wei, L., Lawn, B. R., and Berndt, C. C., 1996, ''Contact Damage in Plasma-Sprayed Alumina-Based Coatings,'' J. Am. Ceram. Soc., **79**, pp. 1907–1914.
[29] Pajares, A., Wei, L., Lawn, B. R., Padture, N. P., and Berndt, C. C., 1996, ''Mechanical Characterization of Plasma Sprayed Ceramic Coatings on Metal Substrates by Contact Testing,'' Mater. Sci. Eng., A, **A208**, pp. 158–165.
[30] Fischer-Cripps, A. C., and Lawn, B. R., 1996, ''Stress Analysis of Contact Deformation in Quasi-Plastic Ceramics,'' J. Am. Ceram. Soc., **79**, pp. 2609–2618.

Effects of Alloy Composition on the Performance of Yttria Stabilized Zirconia—Thermal Barrier Coatings

Josh Kimmel

Zaher Mutasim

William Brentnall

Solar Turbines Incorporated,
San Diego, California 92186-5376

Thermal barrier coatings (TBCs) provide an alloy surface temperature reduction when applied to turbine component surfaces. Thermal barrier coatings can be used as a tool for the designer to augment the power and/or enhance the efficiency of gas turbine engines. TBCs have been used successfully in the aerospace industry for many years, with only limited use for industrial gas turbine applications. Industrial gas turbines operate for substantially longer cycles and time between overhauls, and thus endurance becomes a critical factor. There are many factors that affect the life of a TBC including the composition and microstructure of the base alloy and bond coating. Alloys such as Mar-M 247, CMSX-4, and CMSX-10 are materials used for high temperature turbine environments, and usually require protective and/or thermal barrier coatings for increased performance. Elements such as hafnium, rhenium, and yttrium have shown considerable improvements in the strength of these alloys. However, these elements may result in varying effects on the coatability and environmental performance of these alloys. This paper discusses the effects of these elements on the performance of thermal barrier coatings. [S0742-4795(00)02603-X]

Background

The demand for increased gas turbine efficiency and improved durability, while reducing life cycle cost, has evoked a lot of attention from the gas turbine original equipment manufacturers (OEMs) and end users. While OEMs are continuing their drive to improve engine performance by introducing new concepts to engine designs, there are many limitations and challenges that the engine designer experiences. These include the levels of mechanical and thermal stresses that the turbine components will experience.

The performance and durability of the industrial gas turbines are strongly dependent on the operating conditions and the environment in which they function. Advances in materials have progressed at an increasing rate over the last twenty years and as a result new superalloys have emerged to support the new challenging demands and requirements of gas turbines. The evolution of directionally solidified (DS) and single crystal (SX) alloys has resulted in higher temperature creep and fatigue strengths. These improved properties resulted from the addition of refractory elements such as tungsten and rhenium and the implementation of uni-directional and single crystal manufacturing processes. A metal temperature capability improvement of 25°F is achieved with a change from equiaxed to a directionally solidified CM 247 LC alloy, and an additional 30–40°F could be achieved with the use of a single crystal version of the same alloy [1]. Additional temperature gains are achieved by additions of rhenium. This performance improvement could be achieved, but at the expense of the alloy environmental resistance properties. The use of improved cooling schemes and DS or single crystal alloys to raise the temperature capability, is very close to reaching its limit. Thermal barrier coatings (TBCs) are the next step to achieve further temperature gains, or as an alternative to higher cost DS or single crystal alloys. TBCs are used to reduce the surface metal temperature to workable ranges.

A thermal barrier coating consists of a bond coating applied directly to the alloy, and a ceramic top coat applied to the bond coating. The bond coating is an oxidation resistant coating such as a diffusion platinum aluminide or a MCrAlY coating. The function of the metallic bond coating is to provide oxidation protection to the metallic substrate, minimize the thermal expansion mismatch, and provide strain compliance between the substrate and the ceramic layer [2]. The ceramic layer with low thermal conductivity creates a thermal gradient needed across the TBC system. The ceramic coating most commonly used in industry is a 7–8 wt. percent yttria-stabilized zirconia (YSZ), due to the combination of low thermal conductivity and a thermal expansion close to nickel base superalloys.

There is a tremendous amount of experience using TBCs in aerospace engines [3,4]. Unfortunately, there is limited TBC experience in industrial gas turbine environments [5–7], which show substantially different conditions. Industrial gas turbines (IGT) operate for substantially longer cycles and times between overhauls, and thus endurance becomes a critical factor. Table 1 shows the TBC requirements for aircraft and IGT environments. It is clear that industrial gas turbine users need to gain laboratory and engine experience to determine if TBCs can be used as a prime-reliant coating.

This paper discusses the performance of electron beam-physical

Table 1 Aircraft and IGT gas turbines TBC requirements

Requirement	Aircraft	IGT
Service Time (Hours)	8000	30,000
Number of Cycles	1000 – 8000	10 – 3000
Peak Surface Temperature (°F)	>2300	<2000
Environment	Oxidative / Erosive	Oxidative / Corrosive

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-350. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

vapor deposited (EB-PVD), and air plasma sprayed (APS) TBCs on three nickel base superalloys: equiaxed Mar-M247, single crystal CMSX-4, and the new third-generation single crystal CMSX-10 alloy. CMSX-10 alloy is of great interest to the gas turbine industry because of its improved mechanical properties at high temperatures. The effects of temperature, cycle time, and alloy selection on TBC performance will be discussed along with the associated failure mechanisms.

Processing

Mar-M 247, CMSX-4, and CMSX-10 nickel base superalloy buttons (1″ diameter×0.125″ thick) were coated with EB-PVD and air plasma spray TBC systems. The EB-PVD bond coat consisted of a single-phase platinum aluminide bond coating applied by platinum electroplating followed by above-the-pack aluminizing. The 6–8 wt. percent yttria stabilized zirconia TBC was applied using EB-PVD TBC processing. The APS system consisted of a Co-32Ni-22Cr-12Al-0.75Y-1.5Si bond coating, and a 6–8 wt. percent YSZ-TBC top coating. Processing conditions for the coated specimens included surface preparation by grit blasting, heat treating the bond coating to improve its surface finish and adherence for EB-PVD only, and final heat treatments as required per substrate alloy requirements.

Testing

Coated specimens were evaluated for coating microstructure and composition using a scanning electron microscope (SEM) equipped with energy dispersive X-ray (EDX) analyzer. Thermal cycle testing at 2000°F (100 hour cycles), 2000°F (10 hour cycles), and 1900°F (100 hour cycles) were conducted in a still-air conduction furnace using duplicate specimens. The metallographic preparation procedure included mounting the specimens in epoxy resin containers prior to sectioning to minimize cracking during subsequent sectioning and polishing.

Results

As-Coated

Microstructure. Figure 1 is a scanning electron micrograph that displays the above-the-pack PtAl bond coat and EB-PVD TBC top coat microstructure on CMSX-4 alloy. The platinum aluminide bond coat consists of a single phase, two-layer coating as displayed in Fig. 1, where the outer 35 microns layer consists of a single-phase platinum rich beta NiAl matrix. The second layer is a 35 micron thick refractory enriched diffusion zone that is located at the interface. The coating displays a high content of dark inclusions at the center of the coating. The inclusions were identified as alumina grit particles that were used in specimen preparation. Coating thickness was measured at 70 microns. A thermally grown alpha (α-Al_2O_3) is intentionally grown on the platinum aluminide coating to optimize the bond strength of the TBC system. The thermally grown oxide (TGO) is 1 micron thick, and is shown in Fig. 2. The yttria stabilized zirconia TBC is applied by the EB-PVD process, which exhibits a columnar structure with gaps or porosity perpendicular to the bond coat interface. The ceramic layer grows outward from the TGO layer in needle-like columns. The EB-PVD thickness is 160 microns on the CMSX-4 alloy shown in Fig. 1, and ranged between 170 to 210 microns on the Mar-M 247 alloy.

The microstructure of the EB-PVD TBC coating system on CMSX-10 shows a modified PtAl bond coating, and an identical EB-PVD TBC top coating to the Mar-M 247 and CMSX-4 coating system. The modified platinum aluminide coating exhibits a three layer, two-phase microstructure (Fig. 3). The top layer is $PtAl_2$ phases in a Pt-rich NiAl matrix, the intermediate layer consists mainly of Pt-rich NiAl matrix, and the inner layer is the diffusion zone. The total coating thickness was measured at 70 microns. With this particular coating, no topologically close

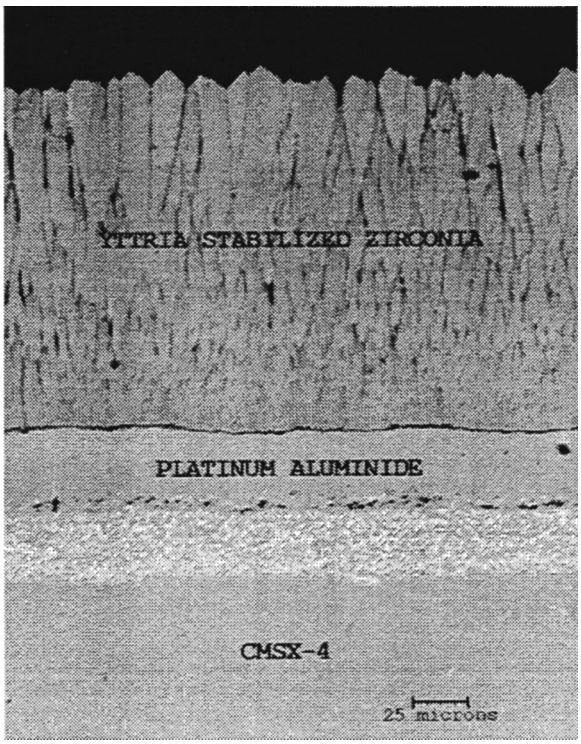


Fig. 1 Typical as-coated thermal barrier coating microstructure of a Platinum Aluminide bond coat and EB-PVD deposited Yttria stabilized Zirconia top coat on CMSX-4 substrate

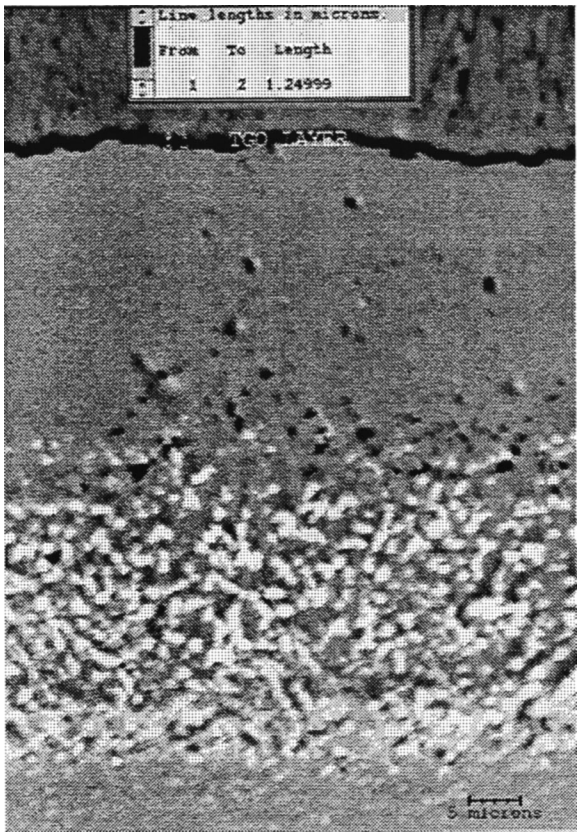


Fig. 2 Aluminum oxide film grown at EB-PVD TBC top coat/ Platinum Aluminide bond coat interface for CMSX-4 specimen

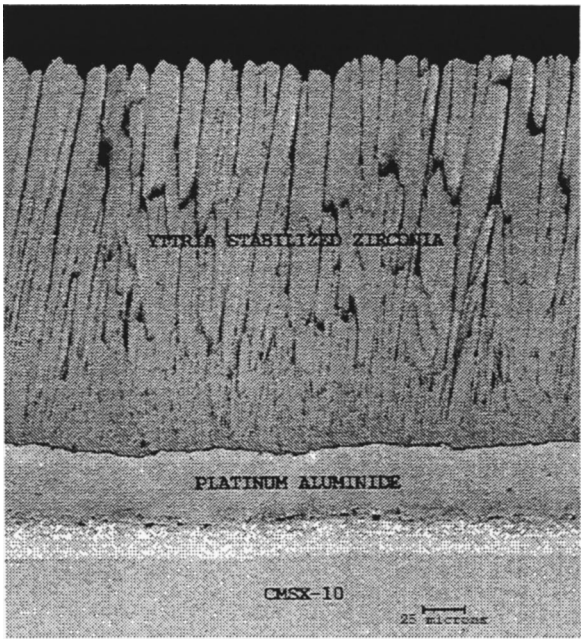


Fig. 3 Typical as-coated thermal barrier coating microstructure of a Platinum Aluminide bond coat and EB-PVD deposited Yttria stabilized Zirconia top coat on CMSX-10 substrate

packed (TCP) phases were present. The EB-PVD thickness is 210 microns on the CMSX-4 alloy, with an identical microstructure to the Mar-M 247 and CMSX-4 coating system.

The APS Co-32Ni-22Cr-12Al-0.75Y-1.5Si bond coating, and YSZ-TBC top coating were processed identically for Mar-M247, CMSX-4 and CMSX-10 alloys. Figure 4 shows the representative coating microstructure on CMSX-4. The thickness of the CoNiCrAlY-Si bond coatings range from 150 to 200 microns. The APS YSZ-TBC top coating typical microstructure displayed in Fig. 4 is formed by building up consecutive layers of ceramic,

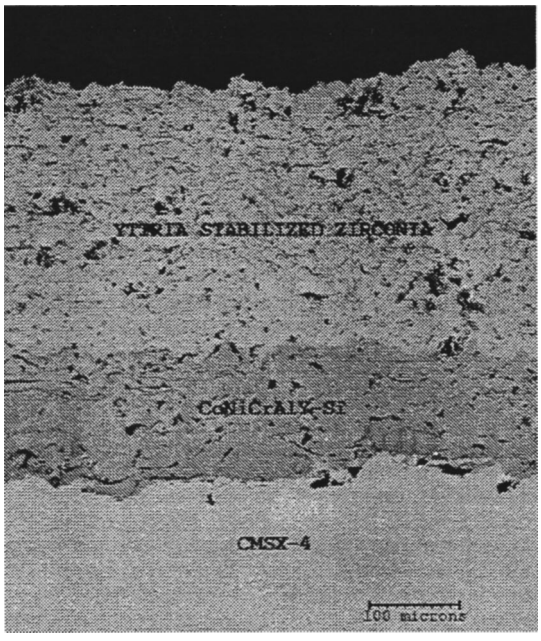


Fig. 4 Typical as-coated thermal barrier coating microstructure of an air plasma sprayed CoNiCrAlY-Si bond coat and Yttria stabilized Zirconia top coat on CMSX-4 substrate

each less than 13 μm thick. The porosity generally appears as equiaxed, round pores that are not linearly connected to one another. The thermal sprayed ceramic top coat thickness ranges from 250 to 300 microns. Thermal spray processes require larger tolerances due to the macroscopic nature of the deposition techniques and rely on rough bond coating surfaces for adequate mechanical adhesion.

Composition. The platinum aluminide bond coatings on Mar-M 247 and CMSX-4 showed some variations, however both coatings exhibited the similar microstructures. The platinum showed 25 wt. percent on the Mar-M 247, and only 13 wt. percent on the CMSX-4. The aluminum showed similar results with 13 wt. percent and 15 wt. percent on the Mar-M 247 and CMSX-4, respectively. The higher amount of platinum in the Mar-M 247 alloy may act to improve the stability of the thermally grown oxide. The aluminum concentration is lower than expected for a beta phase NiAl that is between 18 and 35 wt. percent. However when using the SEM/EDX technique a lower Al wt. percent is usually measured. The platinum aluminide coating on CMSX-10 alloy is a two-phase coating, and therefore exhibits a higher platinum weight percentage. The coating exhibited 32 wt. percent platinum and 25 wt. percent aluminum. The APS Co-32Ni-22Cr-12Al-0.75Y-1.5Si bond coating had aluminum contents ranging from 8–11 wt. percent aluminum at the bond/top coating interface on all alloys. All EB-PVD and APS thermal barrier coatings showed 6–8 wt. percent yttria stabilized zirconia.

2000°F/100 h Thermal Cycles

Visual Examination. Figure 5 displays all TBC specimens after thermal cycling. The thermal cycling test was terminated at coating failure or spallation of coating, with the results displayed in Table 2. Failure was determined at the initial stages of ceramic coating spallation. TBC failure ranged from (1) cycle to those that lasted (51) 100-h cycles. It is interesting to note the obvious failure modes at this point. All the EB-PVD TBC systems showed spallation that initiated at the edges of the specimens, and had different sizes of spallation (Fig. 5 *a, b,* and *c*). The TBC had completely separated, suggesting that the spallation occurred at the bond coat-TGO interface. The APS TBC systems showed complete separation of the ceramic on the entire specimen that initiated at the edge of the specimens. The APS TBC on Mar-M 247 and CMSX-4 failed within the ceramic, just above the bond coating. This is represented by the remaining ceramic on the failed coupon (Fig. 5*e*). The APS TBC on CMSX-10 alloy exhibits a completely blue surface at the debond, representing a failure within the bond coating (Fig. 5*f*). The EB-PVD and APS TBC specimens all showed failure that mitiated at edges due to the coating tensile stresses at the specimen ends. It is evident that TBC life is dependent on many factors including the alloy and bond coating selection, and processing method. These factors will be discussed in detail in the discussion section. After the termination of the cyclic test, the specimens were mounted in epoxy, and sectioned for metallurgical evaluation in order to understand the TBC interfacial behavior after thermal exposure.

Microstructure. The failure of the APS TBC systems occurred within the ceramic approximately 25–50 microns above the bond coating–TBC interface, and the microstructure is shown in Fig. 6. The ceramic layer showed a noticeable amount of sintering. The densification of the ceramic layer due to sintering was accompanied by micro-crack formation within the ceramic layer. The micro-cracks act as stress relief sites within the ceramic layer. Severe lateral cracking is also evident in the ceramic that contributed to the failure of the TBC system. An aluminum oxide phase is formed at the interface of the coating that acts as glue in joining the metallic-ceramic system. A pure alumina phase represented by the dark gray areas, and a spinel represented by a light gray phase

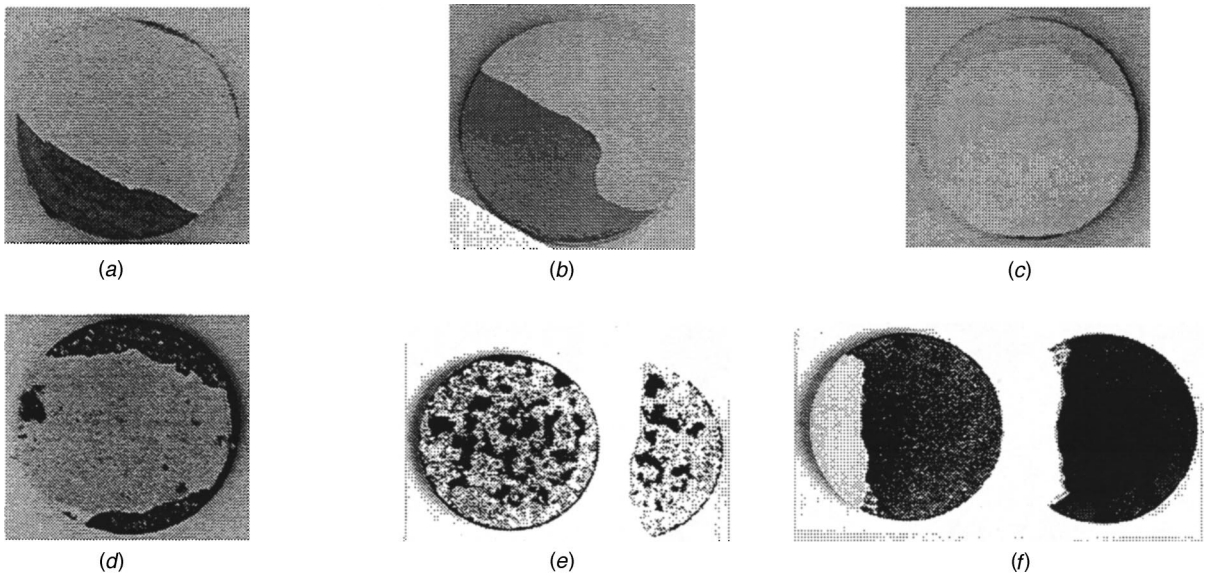


Fig. 5 Thermal cycle oxidation test specimens in post-test condition: (*a*) EB-PVD TBC on MAR-M 247 alloy; (*b*) EB-PVD TBC on CMSX-4 alloy; (*c*) EB-PVD TBC on CMSX-10 alloy; (*d*) APS TBC on MAR-M 247 alloy; (*e*) APS TBC on CMSX-4 alloy; and (*f*) APS TBC on CMSX-10 alloy.

were formed at the interface. The spinel phase is less stable than the α-alumina and can lead to a failure initiation site. There has been a vast amount of study on the mechanisms that cause the failure in the metallic-ceramic bond on plasma sprayed systems [8]. The failures shown here relate to the classical case where tensile stresses are generated at the peaks of a bond coating during cooling to room temperature. Cracks are generated at or above these peaks, and cause coating separation based on a combination of crack joining and growth processes. The samples also showed a small degree of delamination at the alloy bond coating interface due to the oxidation of the alloy from the sides of the sample. This is the major failure mechanism in the APS TBC system on CMSX-10 alloy, as shown in Fig. 7. CMSX-10 the sides of the sample. This is the major failure mechanism in the APS TBC

Table 2 Cycles to failure for APS and EBPVD TBC systems. 2000°F for 100, hour cycles.

SUBSTRATE ALLOY	EB-PVD TBC	APS TBC
Mar-M 247	12	45,51
CMSX-4	1,2	18
CMSX-10	6,7	19,26

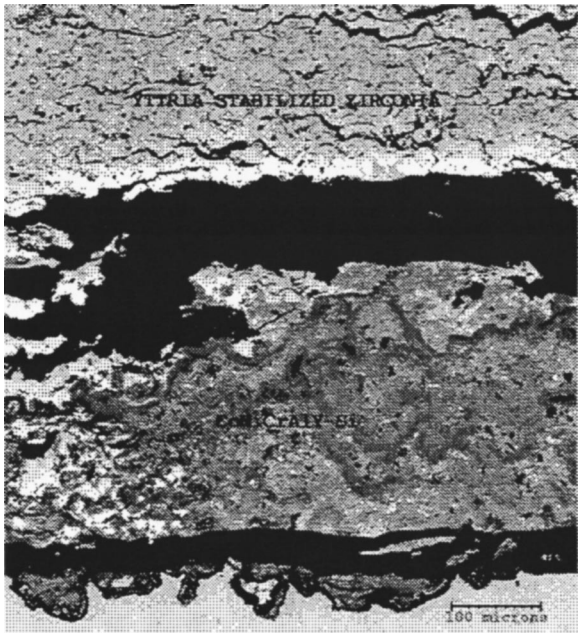


Fig. 6 Air plasma sprayed TBC microstructure on CMSX-4 substrate after thermal cycle testing

Fig. 7 Air plasma sprayed TBC microstructure on CMSX-10 substrate after thermal cycle testing

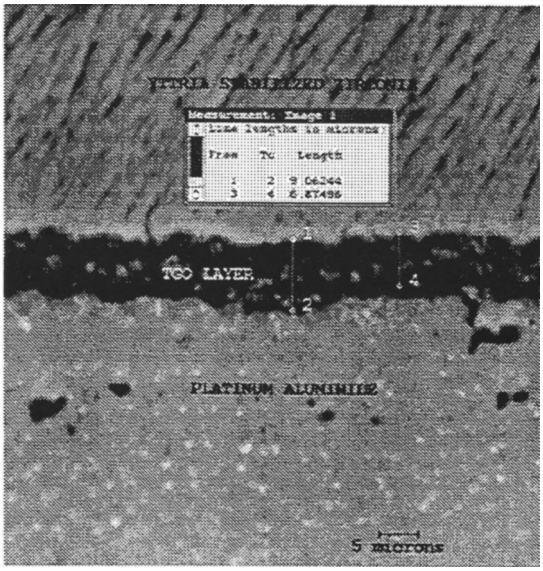


Fig. 8 Thermally grown oxide (TGO) layer at EB-PVD TBC top coat/Platinum Aluminide bond coat interface at the edge of the CMSX-4 specimen after thermal cycle testing

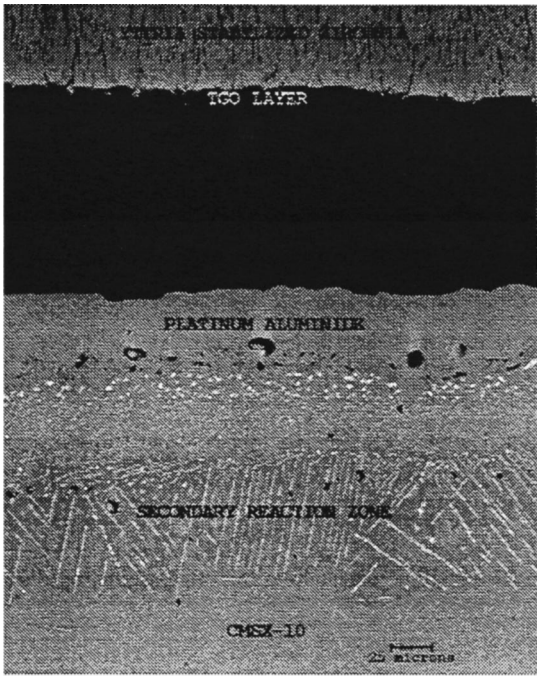


Fig. 10 Cross section of the Platinum Aluminide bond coat and EB-PVD top coat TBC on CMSX-10 substrate after thermal cycle testing

system on CMSX-10 alloy, as shown in Fig. 7. CMSX-10 shows poor oxidation resistance at temperatures above 1800°F [9]. The CMSX-10 alloy formed a thick oxide layer that spalled off the bond coating and TBC. The coating also showed a certain degree of cracking within the ceramic and metallic-ceramic interface. It is believed that the sample was close to failure within the ceramic layer based on the microstructure.

The EB-PVD TBC coatings failed at the platinum aluminide-TGO interface. The failures shown are the classic buckling mechanism, and are caused by the stresses on the TGO layer. The stresses can be caused by factors such as roughness at the TGO, sintering of the YSZ, interfacial impurities, and growth stresses of the TGO [8]. There was no noticeable sintering in the ceramic layer. The EB-PVD ceramic coating on CMSX-4 debonded at the edge of specimen and propagated across the specimen. The thickness and microstructure at the edge and center of the specimen demonstrate this conclusion. The TGO layer at the edge of the specimen (Fig. 8) is a spinel oxide between 7 and 9 microns thick. The combination of the spinel phase that is less stable than the α-alumina, and the high growth stresses at 7–9 microns thickness caused the spallation. The TGO layer in the center of the coating (Fig. 9) is a more stable 3 micron α-alumina layer. It is believed that the low life of the samples (1 or 2 cycles) is related to the processing of the coating. The samples were platinum plated on the ceramic coated side only, leading to excessive oxidation from the sides of the specimen. The CMSX-4 samples will be retested with a platinum aluminide bond coat all over, and the EB-PVD TBC on one side. The EB-PVD TBC samples failed after 6 and 7 cycles, with the microstructure shown in Fig. 10. These samples were processed with the bond coat all over and the ceramic coating on one side only. After failure, the TGO layer consists of a 5 micron α-alumina layer.

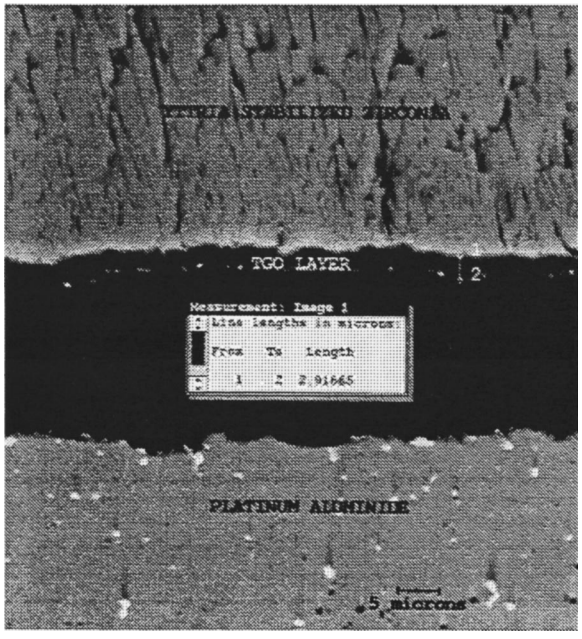


Fig. 9 Thermally grown oxide (TGO) layer at EB-PVD TBC top coat/Platinum Aluminide bond coat interface at the center of the CMSX-4 specimen after thermal cycle testing

Composition. Energy dispersive x-ray analysis was conducted on the oxidized specimens to determine the composition of the various phases in the TBC specimens. The focus was on determining the chemistry of the oxide scales, and the amount of remaining aluminum near the oxide scale. As was clear from the scanning electron micrographs shown in Figs. 6–10, two oxide scales were identified: pure α-alumina and a mixed oxide layer. The mixed oxide scale (spinel) was identified as (Al,Co,Ni)O which has different thermo-mechanical and physical properties than a-alumina scale. Table 3 lists the aluminum concentrations of the TBC systems before and after the thermal cycle test. A remaining aluminum concentration as low as 1 wt. percent was shown for APS TBC system. However, aluminum levels as high as 14 wt. percent Al in the platinum aluminide and TBC system resulted in failure.

Table 3 Aluminum content versus cycles to failure in bond coating

Coating	Substrate	Al wt.% Before Test	Al wt.% After Test	Cycles to Failure
PtAl / EB-PVD TBC	Mar-M 247	13	11	12
PtAl / EB-PVD TBC	CMSX-4	15	14	1,2
PtAl / EB-PVD TBC	CMSX-10	25	13	6,7
APS CoNiCrAlY-Si / TBC	Mar-M 247	8	1	45,51
APS CoNiCrAlY-Si / TBC	CMSX-4	8	1	18
APS CoNiCrAlY-Si / TBC	CMSX-10	9	1	19,26

Discussion

The performance of thermal barrier coatings is dependent on many factors that include the chemistry and microstructure of the bond and top coatings as well as the process with which the coatings are applied. This study focused on the failure modes and the performance of these coatings. These included the effects of the substrate on the lifetime of an APS and EB-PVD TBC system. The selection of the temperature and cycle time on the lifetime is shown to have a substantial effect.

APS Versus EB-PVD Thermal Barrier Coating Systems. Two main top coating microstructures were studied. A plasma sprayed coating with equiaxed grains and circular porosity, and an EB-PVD coating with columnar grains and uni-axial porosity. The EB-PVD and APS TBC specimens all showed failure that initiated at edges due to the coating stresses at the specimen ends. The free edge coating stress field demands that the in-plane thermal mismatch stress reduce to zero at the edge of the ceramic. The transition of the in-plane stress to zero causes strong tensile out-of-plane stress along the coating interface [10]. The coating failure mechanisms and a comparison between lifetimes will be discussed.

TBC spallation is a thermo-mechanical phenomenon that results in the formation of cracks within the TBC layers which propagate with time leading to separation of the ceramic from the metallic layer. Bond coating oxidation and build up of cumulative stresses in the TBC system have been identified as the main events that lead to coating spallation [8]. Two main coating spallation mechanisms were identified in this study. One, which was consistent with the plasma sprayed TBCs, demonstrated failure within the ceramic layer, about 25–50 microns above the oxide/bond coating interface. The second mechanism, which was consistent with the EB-PVD TBCs, showed a planar failure either along or within the oxide layer. These mechanisms are strongly dependent on the type of bonding in each system. The plasma sprayed TBCs depend strongly on the mechanical interlocking between the individual splats in the ceramic coating, and between the ceramic and bond coatings. On the other hand, The EB-PVD TBCs depend on the chemical/metallurgical bonding within the ceramic layer, and between the intentionally thermally grown oxide (TGO) layer and the ceramic top coating. During thermal exposure, a buildup of compressive stresses occurs within the oxide later, and cracks initiate and propagate, and ultimately lead to coating failure. The cracks originate at the weakest location in the TBC system. For plasma spray coatings, this location is usually within the splats near the peaks of the bond coating layer, whereas, it is within the oxide layer of the EB-PVD system.

Table 2 shows a clear improvement in the cycles to failure for the APS system over the EB-PVD TBC systems for different temperatures at 100 hour cycle times. The long cycle times in industrial gas turbines have a detrimental effect on the life of EB-PVD TBC systems. EB-PVD TBC systems that lasted 1000 hours in short burner rig cycle times lasted only 90 hours with 10-hour cycles, and only 40 static isothermal hours in furnace tests [10]. The more porous APS system effectively accommodated the stresses created within the TGO and ceramic layers, and resulted in a 2–3 fold lifetime improvement. In addition, the APS system is more tolerant to a thicker oxide scale than the EB-PVD system, and is stable at much lower levels of aluminum (Table 3). The lower levels of aluminum at failure of the APS TBC systems can be attributed to the presence of yttrium and silicon in the bond coating. Yttrium and silicon are expected to improve the chemical stability of the oxide, which is translsted to an overall improvement in coating life [11]. The higher compliance APS bond coat will not only reduce stresses within the oxide layer, but also reduce thermal stresses within the top ceramic coat, especially when asperities and edges are involved. The EB-PVD system showed only the formation of an α-alumina TGO layer up to 5 microns thick, while the APS system showed the formation of varying levels of spinel oxide formation. The presence of spinels and other mixed oxides, which have lower mechanical strengths than α-alumina, along the bond/top coating interface would result in increasing the mechanical stresses at the interface and in turn would result in coating failure.

Effect of Alloy Composition. Table 2 shows a clear improvement in the cycles to failure for the thermal barrier coatings on Mar-M247 equiaxed alloys versus CMSX-4 and CMSX-10 single crystal alloys. A two-fold improvement was shown for the APS TBCs, and at least a four-fold improvement for the EB-PVD

Table 4 Alloy compositions

Alloy	Ni	Co	Cr	Mo	W	Ta	Nb	Al	Ti	Hf	Re
Mar-M247	bal.	10	8.4	0.7	10	3	0	5.5	1.1	1.4	0
CMSX-4	bal.	9	6.5	0.6	6	6.5	0	5.6	1	0.1	3
CMSX-10	bal.	3	2	0.4	5	8	0.1	5.7	0.2	0.03	6

Table 5 Thermal cycle testing temperature versus cycles to failure

Coating	Substrate	2000°F for 100 Hour Cycles to Failure	1900°F for 100 Hour Cycles to Failure
PtAl / EB-PVD TBC	Mar-M 247	12	>29
PtAl / EB-PVD TBC	CMSX-4	1,2	>19
PtAl / EB-PVD TBC	CMSX-10	6,7	>34
APS CoNiCrAlY-Si / TBC	Mar-M 247	45,51	>29
APS CoNiCrAlY-Si / TBC	CMSX-4	18	>34
APS CoNiCrAlY-Si / TBC	CMSX-10	19,26	>34

Table 6 Thermal cycle testing time versus cycles to failure

Coating	Substrate	2000°F for 100 Hour Cycles to Failure	2000°F for 10 Hour Cycles to Failure
APS CoNiCrAlY-Si	Mar-M 247	45,51	257

TBC coatings. CMSX-10 alloys exhibit a small increase in cycles to failure over the CMSX-4 alloy; however, more testing is needed to verify this conclusion.

The three alloy compositions are shown in Table 4. Compared to the baseline alloy Mar-M 247, reductions in chromium content were needed to accommodate the increasing addition of rhenium in order to maintain alloy microstructural stability. The variations in the alloy chemistries play an important role in the alloy oxidation performance, their coatability and coating behavior. The change from equiaxed to single crystal alloys with high concentrations of refractory elements brings greater challenges to the coatability and oxidation performance of these alloys. For reference, CMSX-10 alloy contains 19.5 wt. percent refractory elements as compared to 16.2 wt. percent for CMSX-4 and 15.1 wt. percent for Mar-M247 alloy. This variation has a positive impact on the alloy strength, but a negative impact on the coatability and performance of these alloys [9]. The decrease in the oxidation performance for the single crystal alloys decreases the stability of the TGO layer, and, therefore, decreasing coating life. In addition, the diffusion of refractory elements such as tantalum (Fig. 9, white particles) to the TGO layer has been shown to act as an initiation site for TBC failure. The higher levels of refractory elements, and lower level of chromium in single crystal alloys exhibits a negative impact in coating lifetime.

Effect of Cyclic Testing Temperature. Thermal cycle testing at 1900°F for 100 h cycles were conducted on Mar-M 247, CMSX-4, and CMSX-10 samples with the same coatings as described above. The results are shown in Table 5, and clearly display an increase in cycles to failure by decreasing the temperature from 2000°F to 1900°F. The lower testing temperature decreases the rate of bond coating oxidation, and therefore lowers the thickness of the oxide layer, thus increasing the cycle to failure. In a similar study, a comparison of coating life at 2100°F and 1900°F indicated approximately 95 percent reduction in life for a 200°F increase in exposure temperature [10].

Effect of Cycle Time. Thermal cycle testing at 2000°F for 10 h cycles was only conducted on the APS TBC system on Mar-M 247 alloy. The results are shown in Table 6 and clearly display an increase in cycles to failure, but a decrease of approximately 50 percent in the total testing time. Mechanical stresses are generated between the metallic and ceramic interface, at or above the TGO layer, during heat-up and cool-down of each cycle due to the thermal expansion mismatch between the various layers. The total cycle time decreased due to increased mechanical stresses on the specimens from cycling. The number of cycles increased because the bond coating oxidation effect is lowered at shorter cycle times.

Conclusions

Thermal barrier coating life is very dependent on many factors that include the chemistry and processing of not only the bond and top coatings, but the chemistry of the substrate alloy as well. APS TBCs exhibited a lifetime improvement over EB-PVD TBCs, and the use of single crystal alloys has a negative effect on the TBC lifetimes. The testing parameters also showed a large effect, with increasing temperatures, or decreasing cycle times showing a negative effect on the lifetimes. Even though the APS system has outperformed the EB-PVD system in thermal cycle testing, the strain tolerance provided by the EB-PVD system must be factored into the selection of a TBC system for an engine environment.

References

[1] Erickson, G. L., 1995, ''A New, Third-Generation, Single-Crystal, Casting Superalloy,'' J. Met., **47**, No. 4, pp. 36–39.

[2] Brindley, W. J., and Miller, R. A., 1989, ''TBCs for Better Engine Efficiency,'' Advanced Materials and Processes, **8**, No. 89, pp. 29–33.

[3] Maricocchi, A., Bartz, A., and Wortman, D., 1995, ''PVD TBC Experience on GE Aircraft Engines,'' Thermal Barrier Coatings Workshop, NASA Conference Publication 3312, pp. 79–90.

[4] Bose, S., and DeMasi, J. T., 1995, ''Thermal Barrier Coating Experience in Gas Turbine Engine at Pratt & Whitney,'' Thermal Barrier Coating Workshop, NASA Conference Publication 3312, pp. 63–78.

[5] Mutasim, Z. Z., Rimlinger, C. M., and Brentnall, W. D., 1997, ''Characterization of Plasma Sprayed and Electron Beam-Physical Vapor Deposited Thermal Barrier Coatings,'' ASME Paper 97-GT-531.

[6] Mutasim, Z. Z., and Brentnall, W. D., 1995, ''Perspective on Thermal Barrier Coatings for Industrial Gas Turbine Applications,'' Thermal Barrier Coatings Workshop, NASA Conference Publication 3312, pp. 103–112.

[7] Nelson, W. A., and Orenstein, R. M., 1995, ''TBC Experience in Land Based Gas Turbines,'' Thermal Barrier Coatings Workshop, NASA Conference Publication 3312, pp. 91–102.
[8] DeMasi-Marcin, J. T., Sheffler, K. D., and Bose, S., 1990, ''Mechanisms of Degradation and Failure in Plasma-Deposited Thermal Barrier Coating,'' ASME J. Eng. Gas Turbines Power, **112**, pp. 521–522.
[9] Mutasim, Z. Z., Kimmel, J., and Brentnall, W. D., 1998, ''Effects of Alloy Composition on the Performance of Diffusion Aluminide Coatings,'' ASME Paper 98-GT-401.
[10] Meier, S. M., Nissley, D. M., and Sheffler, K. D., 1991, ''Thermal Barrier Coating Life Prediction Model Development: Phase II—Final Report,'' NASA Contractor Report 189111, pp. 47–51.
[11] Gupta, D. K., and Duvall, D. S., 1986, ''Overlay Coatings for Superalloys,'' U.S. Patent No. Re.32, 121.

Reductions in Acquisition Costs for State-of-the-Art Fabrication of CFCC Turbine Engine Combustor Liners

Phillip A. Craig
AlliedSignal Composites Inc.,
Newark, DE 19714-9559

David Godfrey
Manufacturing Resources Inc.,
Cleveland, OH 44101

Through work in two U.S. Department of Energy cooperative programs with industry, the confidence level for utilization of ceramic composite SiC/SiC combustor liners has risen from a prototype curiosity to serious planning and addressing of hurdles for commercialization. One of the most significant hurdles for true commercialization is affordability of the ceramic composite combustor liners. To study the manufacturing costs for CFCC liners, a study was conducted to identify cost items and relationships. This report describes the effort, identifies the high impact areas of manufacturing costs, and recommends potential approaches to reducing the manufacturing costs.
[S0742-4795(00)02703-4]

Introduction

The U.S. Department of Energy has two programs that are advancing the technology of future turbine engines, with the goals of reducing U.S. energy consumption and reducing emissions. These programs are CSGT (Ceramic Stationary Gas Turbine) and CFCC (Continuous Fiber Ceramic Composites for industrial applications) [1]. The programs serve U.S. industry by accelerating work and encouraging risk taking through the use of advanced materials, such as ceramic composites instead of metals.

The CFCC Program is a 10-year joint Research and Development effort initiated in 1992 to develop processing methods to produce reliable cost-effective CFCC components for industrial applications. Applied to combustion liners for gas turbines the potential is reduced emissions and increased efficiency. For the DOE to be successful, they must have commercial acceptance, which is a function of customer benefits versus manufacturing costs.

The CSGT Program has the objective to provide an environment for high-risk utilization of ceramics in a stationary power generation turbine engine. This program is being led by the U.S. DOE Office of Industrial Technologies in partnership with Solar Turbines Inc. [2], of San Diego, CA. In this project, a Centaur 10 kW engine, has been upgraded with silicon nitride first stage blades and vanes, and silicon carbide ceramic composite combustor liners. Solar's testing has been widely reported, and includes numerous rig tests, in-house engine tests, and two long duration-field tests with ARCO Energy Company, Bakersfield, CA, of 950 h and 2400 h [3].

From these tests, the confidence level for utilization of the CVI Enhanced SiC/SiC combustor liners has risen from a prototype curiosity to serious planning and addressing of hurdles for commercialization. One of the most significant hurdles for true commercialization is affordability of the ceramic composite combustor liners.

Manufacturing Resources, Inc. (MRI) [4] was contracted through Oak Ridge National Laboratory (ORNL) as a task of the CFCC Program, to study the manufacturing costs for CFCC liners. In the study, they are to identify cost items and relationships, and recommend initiatives to reduce costs to accelerate the commercialization of CFCC technology.

This report describes the MRI effort, identifies the high impact areas of manufacturing costs, and recommends potential approaches to reducing the manufacturing costs. It also addresses issues that influence the choice of advanced materials. Ideally, materials selection for parts, such as combustors, is not just driven by initial acquisition costs but should also consider life cycle costs and time to market. Life and the ability to refurbish the liner for continued service are good examples of key factors in determining an acceptable ceramic composite combustor liner, beyond initial price.

For this analysis, we group all affordability categories into two basic groups; first the value in use including all operational objectives, and second the initial acquisition cost.

Value in use is provided by increased efficiency of combustion, resulting in reduced emissions from combustion. This reduces downstream pollution remediation equipment which may be required to meet regulated levels. Value is also provided, or lost, by life of the liners, in MTBO (Mean Time Between Overhauls) and total life. Value is also provided by increased fuel efficiency resulting in decreased dollars per kilowatt-hour generated. All of these value-in-use-factors have been estimated, and have been reported by Solar Turbines in both open literature and proprietary documents.

The result of this analysis by Solar and other CSGT team members has been estimated for this study. The value-in-use-factors suggest that ceramic combustor initial acquisition prices may be able to be twice the cost of the metal liners (Fig. 1).

The MRI study was given a general cost target for production of about 120 engine sets of CFCC liners per year. MRI's goal was then to determine from process and raw material costs, and a manufacturing scenario where cost targets might be reached.

Previous cost studies from the material supplier suggested that with a wide range of variables, estimates of a production CFCC liner set were 2× to 7× times today's commercial metallic liner. MRI also had the objective to apply greater accuracy to this estimate of combustor pricing.

Manufacturing cost estimates of advanced ceramic composites are complicated due to the immature state of the industry today. Today, two major processes of CFCC manufacture are being utilized for the combustor liners, CVI (Chemical Vapor Infiltration) for SiC/SiC materials and MI (Melt infiltration) for SiC/Si/SiC

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-352. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

ATTRIBUTE	RANGE	COMMENTS
Life	10,000 to 30,000 hours	With less than 10% stress capability reduction
Wall Temperature	2000 F to 2400 F	
Outer Liner Diameter	30" to 45"	Includes other engines
Height	8" to 9"	
Wall Thickness	.1" to .2"	Subject to design changes
Water Vapor	15% at 10 atm.	Nominal
Combined Stress	15 Ksi	

Fig. 1 CFCC combustion liner requirements

materials. In practice, the processes are performed uniquely by different vendors. So MRI's study of similar performing materials from different qualified vendors could have impact depending on that companies' capital equipment. Therefore, the study was performed with knowledge of the processes from one supplier [5] for CVI and MI composites, but the manufacturing plant recommendations are for an ideal ''greenfield'' plant situation; not for our specific facility. This makes the study useful for the CFCC industry. It provides a model for cost reduction conclusions that may be able to be applied in different ways for different vendors. The study provides conclusions of the highest cost fabrication steps in the CVI and MI process, leading to program tasks to work on the most important problems.

The manufacturing processes for both CVI and MI materials begin with continuous fiber preforms made to the net shape of the combustor desired. Interface coatings are applied to the fibers via CVI to provide a weak debond layer between the fiber and matrix. Matrix densification for the CVI materials is provided by several furnace cycles of SiC infiltration, while Si/SiC MI matrices are formed in one vacuum furnace run. The flow charts of the two processes are similar, with similar cycle time. They both utilize similar complexity of equipment. These many consistencies allow comparison of the processes for the purposes of the MRI cost study [5].

Approach

From the MRI perspective, the cost problem was addressed not as a material expert but rather as a production expert. The approach was to document the process flow for liner manufacturing as it exists today and determine the major cost drivers. Then by trading design options and considering process alternatives to accelerate, compress, or change process steps, MRI could synthesize initiatives to reduce costs.

The assessment of CFCC manufacturing began with the development of a set of functional diagrams which describe the present CVI and MI processes for combustion liners. From these functional descriptions and on-site analyses, cost models were developed with a series of bottom-up costs for both processes fabricating 120 liners/year. The initial volume was based on a reasonable demand for the gas turbine supplier's next generation engine. Components of the cost model were broken down in such a way as to highlight some of the manufacturing options for reducing cost. Variations in the model were then made to test sensitivity of the cost to various initiatives. See Fig. 2.

Based on this approach of projecting what a production manufacturing option for CFCC combustion liners might look like, recommendations were developed for future DOE initiatives that would help reduce cost and accelerate the DOE's ultimate objective of lower emission and higher efficiency gas turbines.

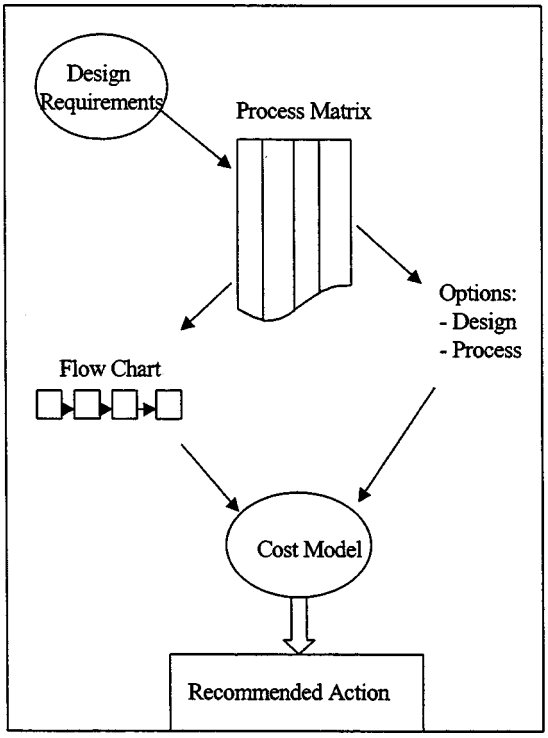


Fig. 2 Functional approach used by MRI. The output from the approach is the table in Fig. 4, showing the relationships between design options, process steps, and costs.

The second portion of the study is a ''should cost'' analysis of the ideal process flow. In this task, a ''greenfield'' process plant is flowcharted, to meet the needs of the production task. Capacity of the plant, future volumes of work, equipment imbalances in ca-

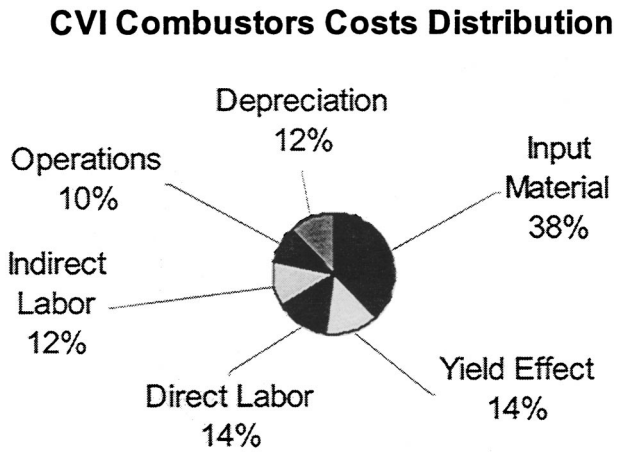


Fig. 3 CVI process costs distribution at 120 combustor sets per year. The ''Input Material'' is predominatly Hi-Nicalon™ fiber. The yield affect includes all losses, also predominatly fiber costs from failed parts. Operations and depreciation combined are about 60 percent CVI reactors, with the rest including building facilities and other equipment. Direct labor is mostly preforming, as a manual process was used in the initial cost model cases. Actual cost estimates in dollars are not shown because of the complexity in representing certain sized parts at certain costs versus other component geometries.

pacity can be studied. Manpower staffing is also studied. No process changes were considered to change the CFCC material.

The study of equipment imbalance is valuable and often the focus of a plant sponsored study on capital spending. The results of the greenfield plant analysis can be used to show the supplier where equipment costs versus two or three shift labor, versus automated equipment seem to be warranted for lowest costs.

The plant analysis by MRI found that a ''greenfield'' CVI plant to produce 120 sets of combustors per year is a feasible operation (Fig. 3). CVI equipment specially sized to fabricate combustors; in arrangements such as placing the smaller inner combustor liner inside the outer liner are the model. The capital costs of the plant consist primarily of the reactors, but by using only the size required for this project, MRI found a large quantity of smaller reactors to be more cost effective at this production rate, and this product mix, than lesser numbers of bigger units. The ''should cost'' analysis is based on a financial layout of the plant breaking even, including all depreciation and amortization costs.

Results Summary

The cost models were run under a variety of assumed operating conditions. The results are summarized below.

CVI and MI Costs are Similar. There is no appreciable difference in cost between the CVI and MI processes for a dedicated facility producing 120 liners/year. While the MI process does require fewer iterations through the CVI reactors or furnaces, this is only a small cost driver compared to the remaining costs. The advantage is off-set by the increased material costs and the additional process steps for MI requiring both labor and support equipment. The initial model for CVI utilizes 100 percent of the reactor capacity at 120 liners/year while the MI uses a smaller portion of the reactor capacity. At greater throughputs the MI process should be matched better with the reactor capacity and gain some additional financial advantage.

Volume Alone Will Not Solve the Cost Problem. At three times the manufacturing volume of the base case, or 360 sets per year, the product cost for the CVI or MI processes is only reduced by 25 percent. If the volume were increased further, the incremental improvement diminishes. Volume by itself will not drive costs down sufficiently to compete with today's liner costs. Liner demand for this class size is projected to be in the 100s, not 1000s.

CVI Reactors Impact the Product Cost Less Than Raw Materials. Any one reactor is estimated at approximately $1 million to build. Amortized over 10 years this means a single reactor represents less than $1,000 to the product cost at 120 liner sets/year. With the present process the total reactor cost impact is significantly less than the raw materials cost. This is an inverse relationship to today's prototyping cost structure.

Fiber Cost is the Leading Cost Driver. The cost of the fiber as purchased input material and as a components lost through less than perfect process yields makes up slightly over 40 percent of the cost. This assumed that at production volumes ceramic fabrics will be obtained at net shape widths to reduce lay-up scrap to zero. Even though the basic processes are relatively reliable by production standards the scrap material costs add up because the raw material cost is such a high multiplier. The cost models assume today's price for Hi-Nicalon fiber is $6,700/kg. The best projected price is about 1/6 of today's price. To reach the lower price the demand for fiber must be greater than 10,000kg/yr. This translates to thousands of gas turbines per year, if combustion liners are the only application and nothing is done to change the fiber manufacturing process costs.

The Next Leading Factor, Direct Labor Including Yield Affects, Represents 17 percent. Preforming is currently the largest use of labor, and automated methods should be investigated, from lessons learned from the organic composites' industry to reduce labor.

Potential Cost Reduction Scenario

Figure 4 is a table showing the relationship between the cost elements, manufacturing drivers and the design features which drive the choice of the cost reduction options. The goal of the chart is to consider design and manufacturing options against the cost elements. This is the output MRI obtains from the functional approach followed via Fig. 2. With the cost models as a guide along with the cost reduction options the following scenario was developed as a test for potential cost improvements.

6:1 Reduction in Material Cost. The cost of Hi-Nicalon flattens out as a function of volume at about a 6:1 reduction in the present procurement price. Therefore, future expectations that this will occur are not unrealistic. The driving force is demand for the fiber. There are two alternate approaches. The first is to find a lower cost material already in demand that meets the present requirements. The problem will be the time it takes to re-certify the material for combustion liner applications.

DESIGN DRIVER	MANU'FNG DRIVER	COST ELEMENT
- Life - Temperature - Wall Thickness - Part Size	- Hi Nicalon Fiber Cost	- Input Material Cost
Part Size	Processing Time for CVI	CVI Process Cost
Thickness	- Process Time - Markets Served - Automation - Process Stability	- Indirect Labor
Part Size	Automation andEquipment	- Direct Labor
	- Miscellaneous Materials - Tooling, Fixtures Life	- Operations Cost
Part Size	-Processing Time - Automation	-Depreciation
	-QC Input Material - Net Shape - Process Stability - Process Control	- Yield

Fig. 4 Cost relationships. Read from right to left. To make an impact on a cost element, consider alternatives in either manufacturing drivers or design drivers. Before changing a design or manufacturing condition, consider the cost element affected.

Reduced Liner Thickness (From 10 Layers to 6 Layers). The next alternative to reduce material cost is to design a thinner liner with fewer layers of material. This will require less material and will impact other costs such as direct labor where lay-up takes less time with fewer layers. The effect of reducing the 10 layers in today's liner to 6 represents a 40 percent reduction in material content and associated labor. Of course, the redesigns would need to consider stiffness and strength issues.

3:1 Increase in Demand. If we now assume a requirement for 360 liner sets instead of 120, is placed in addition to, and on top of the prior improvements, then our product costs are close to half our starting costs. With the increase in demand with this option, additional capital and tooling are required. However, indirect costs are liquidated over a larger set of parts and certain resource utilizations are better matched to the capacity requirements. Actual cost estimates in dollars are not shown, because of the complexity in representing combustors at specific costs versus other geometries of parts.

General Conclusions

The general problem is to get the material application out of the lab and into production. The payoff for any new technology is a function of both costs and benefits. Insufficient attention has been given to quantifying the potential limits for CFCC liners and their competing technologies to define the benefits.

The costs and benefits for CFCC liners must be compared with coated metal liners for the 5PPM-20PPM range emissions and maybe add-on after treatment for even lower emissions if CFCC has the potential to reach that performance. There may be other technologies to consider. The important issue is to require equal attention to benefits as well as costs; therefore, two simultaneous tasks are appropriate:

Task 1. Identify the physical limits and benefit potential for competing technologies.

Task 2. Compare the liner costs with competing technologies.

Product costs can be improved for CFCC liners but not significantly without addressing the input material. If an alternate material for Hi-Nicalon is chosen then time to market becomes a very significant issue. Changing materials seems to be going in the wrong direction to get the liner application out of the lab and into production, unless the time to market can be significantly shortened. Therefore, the following two additional tasks seem most appropriate to address incoming materials costs:

Task 3. Identify the cost drivers for Hi-Nicalon fiber to determine if anything can be done to accelerate cost reduction.

Task 4. Time-line the certification process to identify techniques to accelerate certification of an alternate material.

If the results of the DOE programs continue to support CFCC technology for liners, then the last issue necessary to move CFCC technology from the laboratory environment to a production environment is to create a commercial relationship between the user and the supplier that works. The supplier needs a longer commitment for more than one part. The user needs a proven source he can depend on. The supplier should consider the cost initiatives previously recommended and build a road map to lower costs with measurable milestones at increasing part volumes. The user must be willing to commit to longer-term relationships with larger orders as the supplier meets those milestones. This will be the supplier's incentive to continue to improve. The DOE can supply some initial monetary incentives to mitigate risk and accelerate the process but ultimately the relationship must be commercially viable.

Summary

An outside manufacturing study has been performed that confirms the manufacturing process and plant description for CFCC combustor liners at 2× the acquisition price of today's metal liners. Conclusions from the study are being utilized to plan near term projects to implement the cost reductions.

References

[1] U.S. Department of Energy, 1999, Office of Industrial Technologies, Washington, D.C., Ms. Debbie Haught (202) 586-2211 or Ms. Patricia Hoffman (202) 586-6074, or www.ms.ornl.gov/cfcc/home.htm

[2] Solar Turbines Inc., 1999, a division of Caterpillar, Inc., Mr. Jeff Price, CSGT Program Manager (619) 544-5538.

[3] Van Roode, M., Brentnall, W. D., Smith, K. O., Edwards, B. D., McClain, J., and Price, J. R., 1997, ''Ceramic Stationary Gas Turbine Development—Fourth Annual Summary,'' ASME Paper 97-GT-317.

[4] CFCC Combustion Liner Cost Study, 1998, internal report to ORNL, Manufacturing Resources Inc., Cleveland, OH (Mr. David Godfrey; President, (216) 786-3390).

[5] AlliedSignal Composites, 1999, (formally DuPont Lanxide Composites), Newark, DE 19714. Website: www.dlcomposites.com (Mr. Andrew Miller, webmaster).

An Experimental and Modeling Study of Humid Air Premixed Flames

Anuj Bhargava
e-mail: bhargaa@utrc.utc.com

Med Colket

William Sowa

United Technologies Research Center,
East Hartford, CT 06070

Kent Casleton

Dan Maloney

U.S. Department of Energy,
Federal Energy Technology Center,
Morgantown, WV 26507

An experimental and modeling study has been performed jointly by UTRC and DOE-FETC to determine the effect of humidity in the combustion air on emissions and stability limits of gas turbine premixed flames. This study focuses on developing gas turbine combustor design criteria for the Humid Air Turbine (HAT) cycle. The experiments were conducted at different moisture levels (0 percent, 5 percent, 10 percent, and 15 percent by mass in the air), at a total pressure of 200 psi, pilot levels (0 percent, 1 percent, 3 percent, and 5 percent total fuel), and equivalence ratio (0.4 to 0.8 depending on the moisture levels). The moisture levels were achieved by injecting steam into dry air well upstream of the fuel-air premixing nozzle. Computations were made for comparison to the experiments using GRI Mech 2.11 kinetics and thermodynamic database for modeling the flame chemistry. A Perfectly Stirred Reactor (PSR) network code was used to create a network of PSRs to simulate the flame. Excellent agreement between the measured and modeled NO_x (5–10 percent) was obtained. Trends of added moisture reducing NO_x and the effects of equivalence ratio and piloting level were well predicted. The CO predictions were higher by about 30–50 percent. The CO discrepancies are attributed to in-probe oxidation. The agreement between the data and model predictions over a wide range of conditions indicate the consistency and reliability of the measured data and usefulness of the modeling approach. An analysis of NO_x formation revealed that at constant equilibrium temperature, T_{eq}, the presence of steam leads to lower O-atom concentration which reduces ''Zeldovich and N_2O'' NO_x while higher OH-atom concentration reduces ''Fenimore'' NO_x. [S0742-4795(00)00703-1]

Introduction

The Advanced Turbine System (ATS) Program sponsored by the Department of Energy (DOE) has among its objectives the development of gas turbine-based power systems which reduce the cost of generating electricity while increasing gas turbine efficiency and lowering emissions.

One engine cycle under study is the advanced-cycle Humid Air Turbine (HAT) [1]. The definition and preliminary engineering of a HAT power cycle based on aeroderivative engine technology was undertaken by a United Technology Corporation team, Fluor Daniel, Inc., Texaco, and EPRI in 1991 [2]. This cycle, shown simplistically in Fig. 1, uses low temperature gas turbine exhaust heat as well as heat from within the cycle (intercooler and aftercooler) to supply heated water to a saturator. Here, the compressor discharge is humidified and then the water-vapor laden air passes through a recuperator and combustor and then to the turbine. Both the increased mass flow of the humid air and higher specific enthalpy of the mixture provide additional power but require no additional compression work from the turbine, resulting in efficiency and power output gains.

Several preliminary design studies have been performed on the HAT cycle [3]. HAT cycles were identified that had over 60 percent efficiency and emissions estimates that met the DOE goals. These emissions, however, were based on estimates of NO_x formation in the presence of water vapor. A series of experimental and modeling studies, described in the following sections, have been performed to characterize the effect of humidity on the emissions and combustion stability of HAT cycle configurations.

Several studies have been undertaken to understand the influence of moisture in air on NO_x and CO emissions. Dryer [4], Miyauchi et al. [5], Touchton [6], Blevins and Roby [7], and Meyer and Grienche [8] have all investigated the influence of moisture on emissions under different operating conditions. Most of these studies have dealt with diffusion flames. In diffusion flames, it was concluded that almost all of the NO_x reduction took

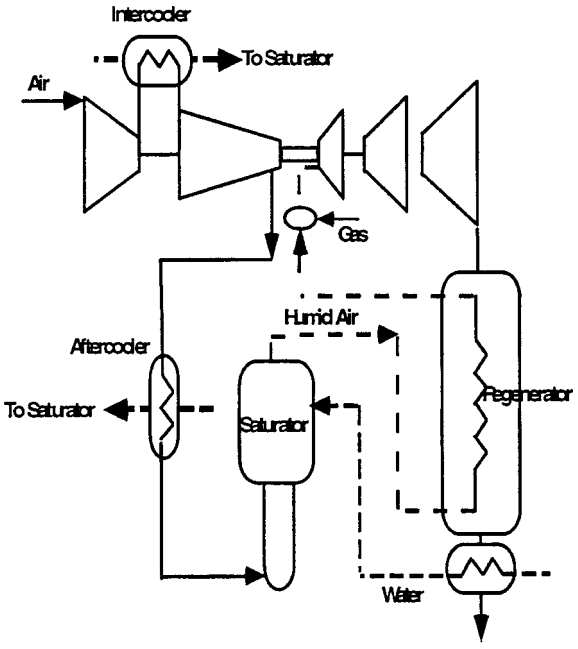


Fig. 1 A schematic of the HAT cycle

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-8. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

place due to a reduction in the flame temperature. In a diffusion flame, where the overall equivalence ratio may be as low as 0.4–0.5 the high temperature stoichiometric regions are the primary source of NO_x. Moisture reduces the peak temperatures which brings about a drastic reduction in NO_x emissions. Not much work has been carried out for premixed systems, required to meet low emissions regulation. Moisture addition into the air reduces NO_x emissions even in premixed systems. Initial investigation of premixed systems revealed several different pathways, which may account for the drastic reduction in NO_x emissions when moisture is increased. Some of the possible routes include:

- Reduction in equilibrium temperature (T_{eq}).
- Reduction of peak temperature of pilot diffusion flames.
- Equivalence ratio change. (In humid air flames for constant T_{eq}, as the moisture level increases, the equivalence ratio must increase resulting in a lower O-atom concentration.)
- Dilution effect. (In humid air flames for constant T_{eq}, as the moisture level increases, the O-atom and N_2 mole fraction ratios decrease.)
- O-atom suppression. (Presence of moisture promotes the reaction $O+H_2O \rightarrow OH+OH$, leading to O-atom suppression reducing NO_x through the ''Zeldovich and N_2O pathways.'')
- Heat release rate reduction lowering peak temperatures and super-equilibrium O-atom.

Experiments

Experiments were conducted in the Low Emissions Combustor Test and Research (LECTR) facility at the U.S. Department of Energy's Federal Energy Technology Center. The LECTR facility, Fig. 2, is a refractory lined pressure vessel containing test modules configured to meet specific test requirements. The facility was designed for pressurized operations up to 450 psig with access to a preheated (1000°F max) air supply capable of delivering up to 3.2 pps of air.

In the present test configuration, superheated steam was added to the preheated combustion air and then injected into an inlet plenum. The flow was straightened through a series of perforated metal plates and passed through the test nozzle where it was mixed with the primary fuel supply. Two diffusion flame pilot fuel streams were also employed. A side-wall pilot fuel flow was injected through the wall of the combustor approximately one half inch above the tip of the nozzle. In addition, center-body pilot fuel was injected along the centerline of the nozzle. Combustion products passed upward across an area-averaged sample probe which spanned the combustor diameter. The flow continued upward into a plenum where it turned 90 deg before entering a water quench zone where the gas temperature was reduced prior to passing through a pressure control valve and out an exhaust stack.

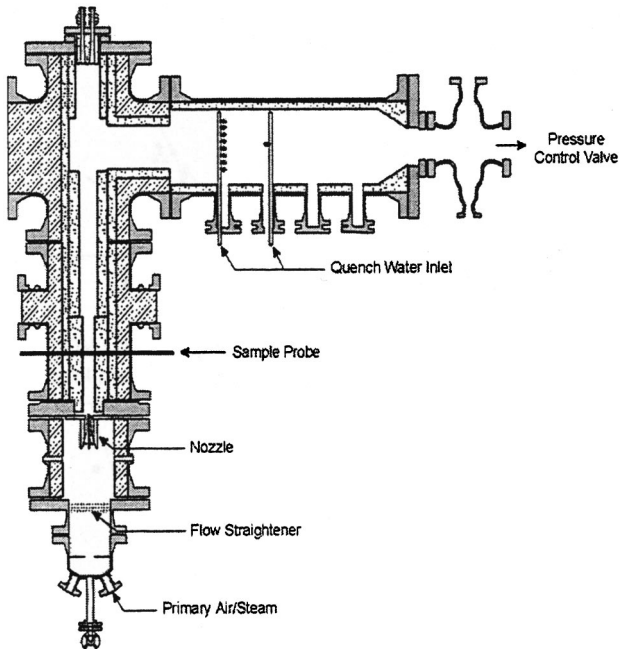


Fig. 2 A schematic of the experimental setup

Experiments were conducted to evaluate the emissions performance as a function of firing conditions for a two square inch effective flow area scaled version of the tangential entry fuel nozzle. Larger scale versions of the nozzle have been reported in the literature [9]. The two square inch effective area nozzle tests were conducted at 200 psi at an air flow of 2.2 pps and a combustor insert diameter of 4.25 in.

Test parameters examined included percentage of total fuel flow as side-wall pilot (0, 1, or 5 percent), percentage of total fuel flow as center-body pilot (0, 1, 3, or 5 percent), percentage of steam relative to combustion air (0, 5, 10, 15, or 20 percent) and air/steam mixture temperature (varied from 935 to 725°F depending on steam loading and air preheat). All percentages were taken as mass percent. All fuel, air and steam flows were measured using precision orifice runs with flow control valves tied to an automated control system. Excellent mass balance closure (typically within 1 to 2 percent) was observed when comparing input flow measurements with measured combustion product compositions.

Experiments were performed by establishing the required air and steam flows for a given test condition. The side-wall, center-body and main fuel flows were then set based on a prescribed equivalence ratio. The nozzle was held at the above condition for a period of 10 to 15 min to establish a steady baseline. The three fuel streams were then simultaneously ramped to lower flows (leaner conditions) keeping their relative ratios constant while the air and steam flows remained fixed. Fuel ramp rates were established based on the gas sample system response time and typically involved reducing fuel flow approximately 30 percent during a one-hour ramp. The starting equivalence ratios for these tests were selected based on previous experience with the test nozzle to give a starting CO emission level of 50 ppm or greater. As the fuel was ramped to leaner conditions, CO emissions passed through a minimum and rose rapidly as the lean blow-off limit was approached. Fuel ramps were discontinued when CO levels exceeded 200 ppm. As steam loading in the combustion air was increased the starting equivalence ratio was increased as well to stay on the rich side of the CO minimum.

Gas samples were obtained using water cooled rake probes designed to extract representative gas samples across the area of the combustor. The sample probe was located approximately 15 in above the nozzle tip. Extracted gas samples were transported through heated transfer lines to a pressure let-down station followed by a chiller to remove water. The dried sample gas was then passed through an analyzer train configured to measure O_2, CO_2, CO, total NO_x, and Unburned Hydrocarbons.

Modeling

Experimental data covered a wide range of equivalence ratios and humidity levels. The fuel concentration in the pilots was also varied from 0 percent to 5 percent fuel in the pilot to determine the optimum level of side pilot required to keep NO_x emissions low, but provide enough stability to the lean flame. Some of the observations made from the experiments were as follows:

For dry flames at constant T_{eq}:

- NO_x emissions increase with increasing fuel in the pilot while CO emissions show no dependence on pilot level

For humid air experiments:

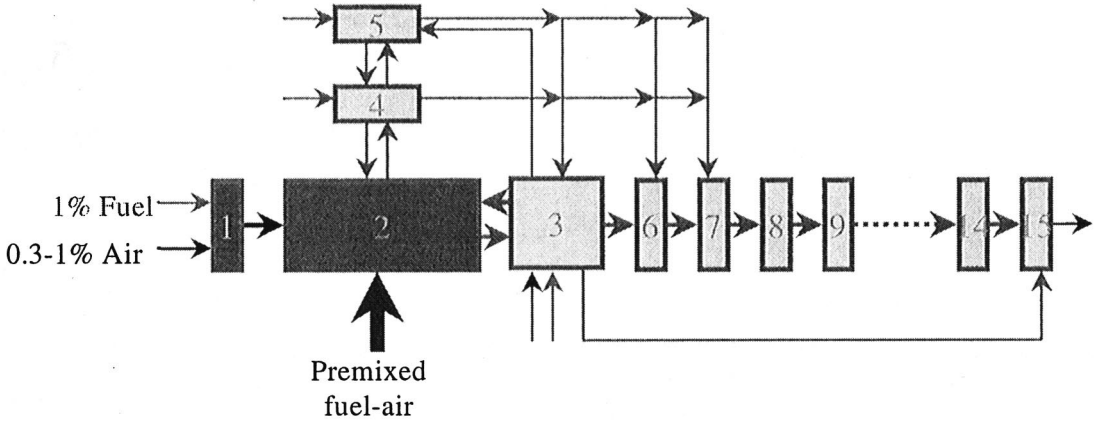


Fig. 3 A schematic of the PSR reactor network used to model the flame

- NO_x and CO levels decrease with increase in the humidity levels
- similar effect of pilot levels as in dry flames

The experiments were modeled using a literature reaction set, GRI Mech2.11 (Bowman et al. [10]) and its accompanying thermochemistry. GRI Mech2.11 reaction set consists of 49 species and 277 reactions and has been validated over a wide range of pressure and equivalence ratios. A PSR Network code developed at UTRC was used to simulate the combustor. The basis for the Network PSR code is the Chemkin PSR model [11]. The Network used to model the experiment is shown in Fig. 3. The network was designed to simulate the mixing and flow characteristics of the experiment, as closely as possible. A total of 15 reactors were used to represent the combustor. The number 15 was arbitrarily chosen to simulate the process in a reasonable manner without a high computational burden. The combined volume of the 15 reactors gave a residence time equal to the experimental residence time in the equipment. A small amount of fuel and air mixture was introduced in the first reactor. This represented the central pilot. The products of reactor 1 along with most of the fuel (75 to 79 percent) and air (80 to 90 percent) were introduced in reactor 2. Twenty percent of the fuel along with 10 percent of air enters the system through reactor 3. The combination of reactors 2 and 3 represent the inhomogeneity that may be present due to less than perfect mixing. Reactors 2 and 3 represent the flame zone. The side pilot was simulated by flowing 0–5 percent of the fuel through reactors 4 and 5. Providing recycle between reactors 2 and 4 simulated recirculation. The rest of the reactors (6 to 15) represent the post-flame zone. NO_x is primarily formed in reactors 2 and 3 while the rest of the reactors, 4 to 15, are for mixing of the side and central pilots with the flame and for CO burnout. In this manner the network was designed to match experimental NO_x and CO. Residence time varied with flow rates as air was fixed while the fuel flow rate was changed in order to achieve the required equivalence ratio. Residence time in each reactor was a function of the flow rate through the reactor and reactor volume. The residence time in all the reactors combined was between 4.5 and 6 ms. The flame zone (reactor 2 and 3) accounted for most of the residence time (nearly 60 percent). The network structure remained the same for all conditions but minor alterations for fuel/air distribution had to be done from one case to another to stabilize the network and to simulate the flow characteristics. For example, with very fuel lean conditions or for cases with 10–15 percent moisture, fuel in the first reactor had to be increased to stabilize the network. With different fuel flows across the pilot, fuel and/or air distribution in reactor 2 and 3 had to be slightly modified.

Solutions were compared with experiments on the basis of combustor exit temperature which is equal to T_{eq} based on local equivalence ratio. The combustor exit temperature in general is different from the local temperature in the PSR as the local temperature can be kinetically limited.

Results and Discussion

Numerous experiments were performed to ascertain the effect of moisture, equivalence ratio, and inlet temperature on the amount of NO_x and CO formed. The effect of pilot level (0, 1, 3, 5 percent fuel) and moisture level (0, 5, 10, 15 percent water in the air by mass) on NO_x and CO emissions can be seen in Figs. 4 to 7.

As seen in Fig. 4, higher side pilot levels lead to higher NO_x. The effect is more pronounced at very lean conditions. This result agrees well with the findings of Leonard and Stegmaier [12], and Maghon et al. [13]. Further analysis of the NO_x formed at different pilot levels in this study indicated that with one percent increase in the fuel injected through the side pilot, NO_x emissions go up by 1.5 ± 0.2 ppm. This increase is less than the increase observed by Maghon et al. [13]. CO emissions remain more or less unaffected by the pilot level, as observed in Fig. 5. The effect of moisture level in the air is seen in Figs. 6 and 7. NO_x and CO emissions show a significant drop with increasing amounts of moisture in the air stream. Much of the reduction shown in these figures can be attributed to a reduction in flame temperatures at constant ϕ, with increasing moisture. As will be shown later, even at constant flame temperature, NO_x and CO is reduced with increasing moisture levels.

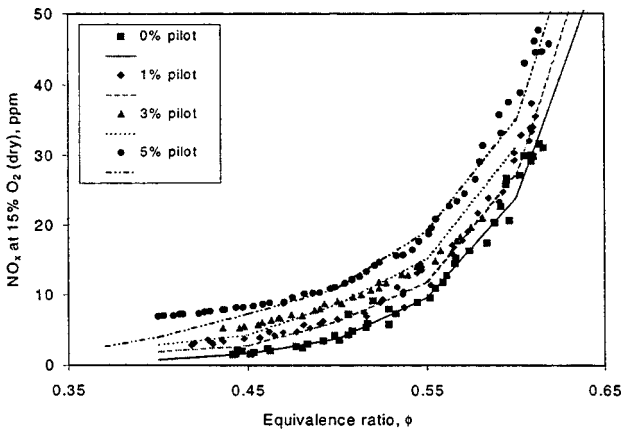


Fig. 4 Comparison between measured NO_x and computed NO_x at 200 psi at different pilot levels

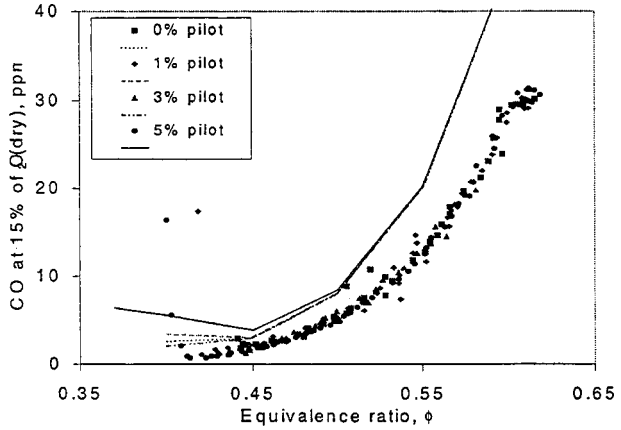


Fig. 5 Comparison between measured CO and computed CO at 200 psi at different pilot levels

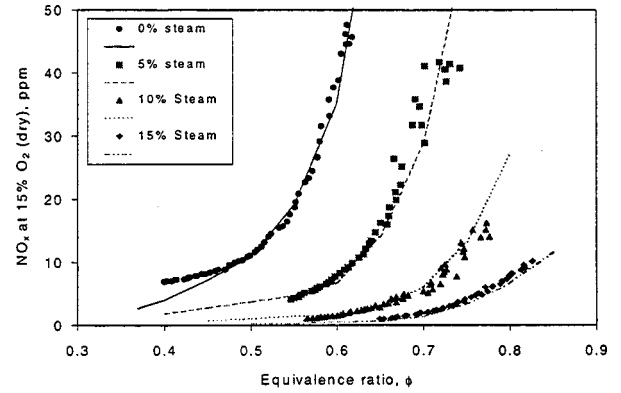


Fig. 6 Comparison between measured and computed NO_x for a 5 percent side pilot flame at 200 psi for different steam loading

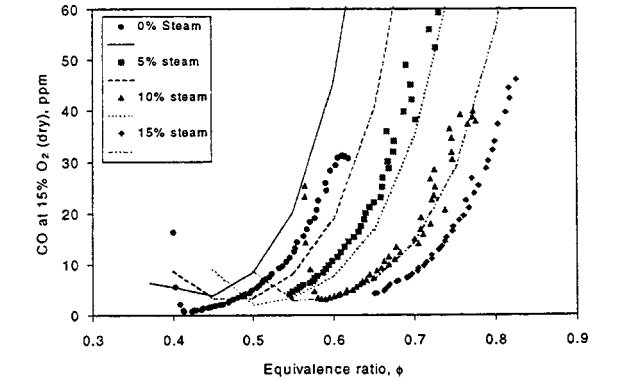


Fig. 7 Comparison between measured and computed CO for a 5 percent side pilot flame at 200 psi for different steam loading

Simulations examined the influence of pilot and humidity level consistent with the experiments. Figs. 4 and 5 show the comparison between the experiments and modeling as piloting level varies. NO_x predictions are very good but the predictions for CO are off by as much as 50 percent. This trend was observed in all simulations. The CO predictions were close to the equilibrium levels and are highly sensitive to the mixture (air and fuel) temperature. A small (~50°F) discrepancy in the measured air-fuel temperature can account for the difference in the measured and computed CO. Alternatively, the sample probe's impact on the equilibrating CO which is still undergoing reaction may explain the lower experimental values. Small levels of additional CO destruction in the probe could also explain the discrepancy. A study of the probe effect on CO measurements has been presented by Nguyen et al. [14]. Nguyen et al. measured CO in an atmospheric pressure flame using line-of-sight tunable diode laser absorption and extractive probe sampling and compared it with CO prediction using a numerical model. They found that the measurements made with extractive sampling technique were as much as 10 times less than the laser-based experiments for fuel-lean flames whereas for fuel-rich flames the agreement between the two techniques was good.

The same network design developed for the dry air cases was extended to humid air cases. Shown in Figs. 6 and 7 are the comparison for NO_x and CO mole fraction measured in the flame and simulated by the network. Higher humidity brings about a drastic reduction in the formation of NO_x. As the amount of moisture increases, the lean blowout limit moves to higher equivalence ratios consistent with the shift seen in Figs. 6 and 7. For a dry air flame at $T_{air}=860°F$ and at $P=200$ psi, the boundary between equilibrium CO and the onset of lean blow out dynamics occurs at an equivalence ratio of 0.41. This boundary shifts as more moisture is introduced. It goes up to 0.55 for 5 percent moisture and to 0.65 for 15 percent moisture in the air stream. The network PSR is able to capture the behavior of the NO_x concentrations to within 5–10 percent of the measurements, further corroborating the modeling approach. CO predictions are higher (20–40 percent) than the measurements. This effect was similar to the one seen in the piloting data.

As seen in Fig. 7, flame stability limits move to higher equivalence ratios as moisture content increases. Moisture in the air reduces flame temperatures. Flames with different humidity levels have different temperatures for the same equivalence ratio. For example, a 0.54 equivalence ratio flame with no steam has the same T_{eq} as a flame with 15 percent steam in the air at an equivalence ratio of 0.75. Another method, therefore, of comparing flames is to compare NO_x and CO emissions formed in flames at different moisture levels as a function of T_{eq}. The computations of T_{eq} were carried out using equilibrium code called STANJAN, developed by Reynolds [15] and modified by Kee and Lutz [16]. The computation included 49 species. The GRI 2.11 thermodynamic database was used for these computations. The temperatures in Figs. 8 and 9 (and also Figs. 10 and 11) are the T_{eq} for the entire fuel-air mixture in the combustor. Actual flame temperatures in the experiment are distributed and somewhat lower due to heat losses to the burner and surrounding air in the chamber. Figures 8 and 9 indicate the dependence of NO_x and CO on T_{eq}. As seen in Fig. 8, calculations and measurements indicate that a

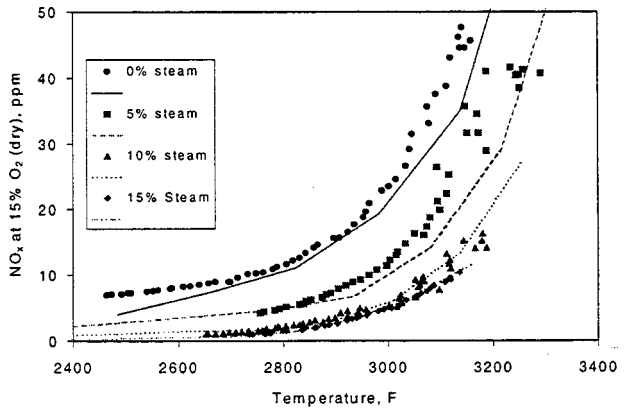


Fig. 8 Comparison between measured and computed NO_x for a 5 percent side pilot flame at 200 psi for different steam loading

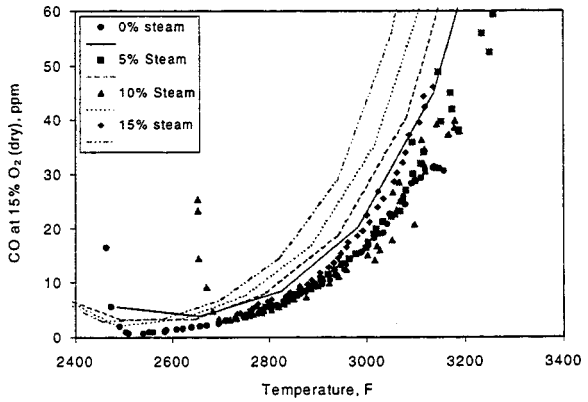


Fig. 9 Comparison between measured and computed CO for a 5 percent side pilot flame at 200 psi for different steam loading

dry air flame at $\phi=0.54$ would have a T_{eq} of 2950°F and NO_x emissions of around 15 ppm. At the same time, a flame with humid air (15 percent steam content) at $\phi=0.75$ has the same T_{eq} of 2950°F, but its NO_x emissions are only 3 ppm. This means that besides reducing NO formation in flames by reducing the flame temperature [6,7], humidity also reduces NO by some other mechanism. This reduction can be due to the change in the radical pool because of varying amounts of moisture and different equivalence ratios at the same temperature, or due to any of the other factors mentioned in the introduction. The flame at 0.54 equivalence ratio (dry air) is richer in O-atom than a humid air flame (15 percent steam) at 0.75 equivalence ratio, even though both flames are at the same T_{eq}. One of the other possibilities which can account for the drastic reduction in NO_x is the temperature attained in each of the PSR's in the network. As mentioned in the introduction, moisture in the feed stream can slow heat transfer rates, and, hence, a different moisture content can lead to a different maximum temperature (residence time in the different

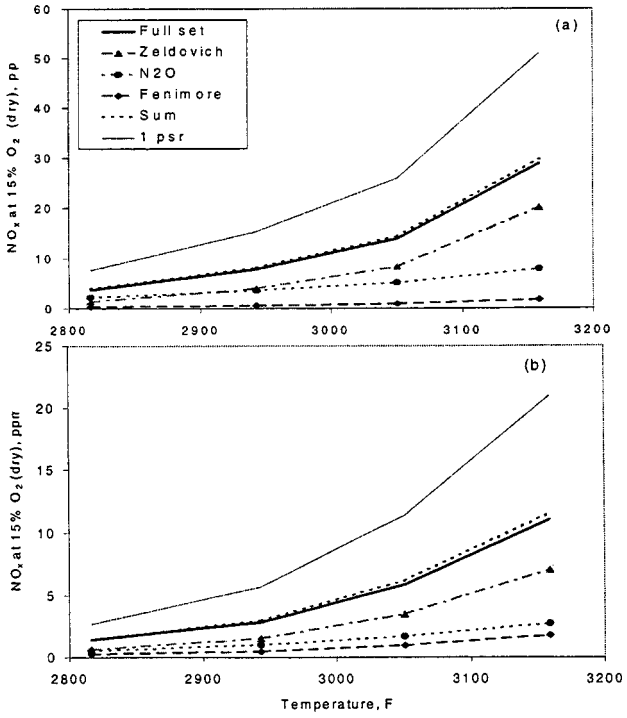


Fig. 10 Influence of different pathways on NO_x formation for a flame with a 0 percent side pilot and (*a*) 0 percent steam and (*b*) 15 percent steam. Calculations with one psr also shown for comparison purposes.

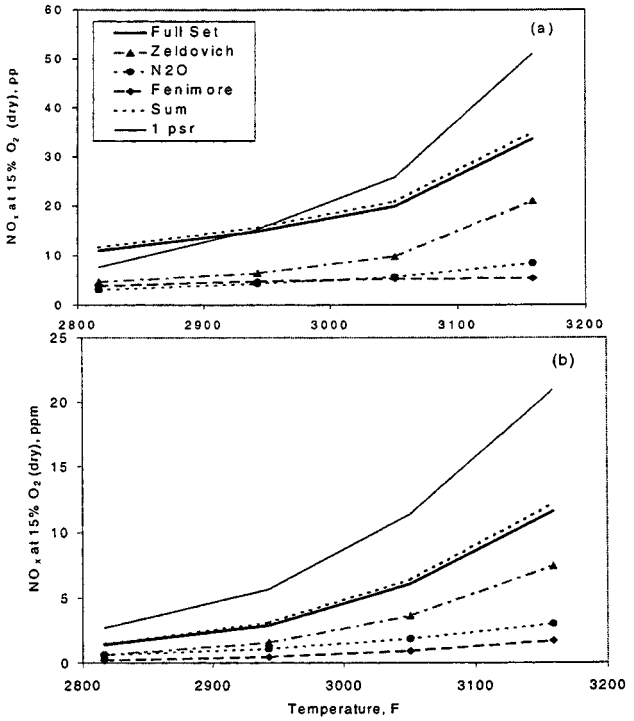


Fig. 11 Influence of different pathways on NO_x formation for a flame stabilized by a 5 percent side pilot and (*a*) 0 percent and (*b*) 15 percent steam. Calculations with one psr also shown for comparison purposes.

PSR's is fixed) attained in PSR 2 and PSR 3 of the network, which account for most of the NO_x formed. An analysis of the temperatures attained in the different PSR's in the network, for flames with different moisture content but at the same T_{eq}, indicated that moisture content does not alter the temperature attained in each of the reactors. Hence, a reduction in NO_x concentration with increased moisture content is not due to lowering of peak temperatures in the reactors. CO emission levels at one T_{eq} do not vary with humidity, as seen in Fig. 9.

To determine the roles of various mechanisms on the production of NO_x, several modifications were made to the reaction set. NO_x formation reactions were divided into three different but coupled groups as follows:

1 ''Zeldovich or thermal NO_x Mechanism'' [17] as extended by Bowman and Seery [18]. NO_x in Zeldovich or thermal NO_x mechanism is initiated by the reaction of N_2 with O. This well known three-step reaction sequence is the dominant NO_x forming route at temperatures above 2950°F. Zeldovich or thermal NO_x mechanism was suppressed by making the reaction $N+NO\rightarrow N_2+O$ proceed in the forward direction only.

2 ''N_2O Mechanism'' consisting of reactions involving N_2O which form NO. NO_x formed via this mechanism can be effectively suppressed by making the $N_2O(+M)\rightarrow N_2+O(+M)$ proceed in the forward direction only.

3 The third mechanism responsible for NO_x formation is ''Fenimore NO'' [19]. All NO_x forming reactions, beside the ''Zeldovich and N_2O'' pathways, are included in this route. The main steps influencing this reaction sequence are $N_2+CH=HCN+N$, $N_2+CH_2=HCN+NH$, and other reactions involving N_2 and hydrocarbon radicals. This extensive sequence was suppressed by inhibiting the forward rates for reactions between hydrocarbon fragments and molecular nitrogen.

The NO_x formation rates of each of these reaction sequences can be enhanced through super-equilibrium levels of O-atoms and other radical species in the flame front.

To study the contribution of each pathway, the other two channels were suppressed. Hence, in order to analyze the importance of ''Zeldovich Mechanism,'' NO_x formation via ''Fenimore and N_2O Mechanisms'' was removed. Similar computations were performed to evaluate the importance of ''Fenimore and N_2O Mechanisms.'' The analysis was carried out for 0 percent and 5 percent pilot levels at 0 percent and 15 percent moisture loading. Figures 10 and 11 show the contributions of the different mechanisms for the four cases. Also shown in the figures is the NO_x computed with a single PSR. The differences between these calculations and those of the network analysis indicates the effect of mixing in the combustion chamber and circulation on NO_x formation and finite reaction times required to approach T_{eq}. The NO_x prediction for the 1 PSR case is higher than the Network PSR, primarily due to residence time distribution. In the Network, reactors 2 and 3 are the reactors primarily responsible for making NO_x and CO. In order to simulate the experiments with the network, reactor volume (which determines the residence time in the reactor) was used as one of the adjustable parameters. Once established for one case, the volumes were not changed for any of the other cases. Hence, by keeping the residence time low, one may inhibit NO_x generation. On the other hand, in the case of 1 PSR, the residence time was not a variable and was the same as the whole of the network combined.

''Zeldovich NO_x'' dominated the other NO_x formation channels in the temperature range studied, except for the 0 percent side pilot flame for dry air. For this case, seen in Fig. 10(*a*), ''N_2O Pathway'' dominates the ''Zeldovich Pathway'' in the lean operating range (2800 to 2900°F). This is the range in which most of the industrial gas-turbines operate.

Figure 10 shows the behavior of NO_x formed at two moisture levels, 0 and 15 percent, for flames with 0 percent fuel coming in from the side pilot. As seen in the figure, at temperatures around 2950°F, NO_x formed by the N_2O pathway is comparable to the ''Zeldovich or thermal'' NO_x, but at temperatures around 3050°F, ''Zeldovich'' NO_x starts increasing much more rapidly and becomes the dominant NO_x formation channel. Comparing the NO_x formed in the flames with and without moisture, it can be seen that at the same temperature, NO_x formed in the cases with 15 percent moisture is much less than the NO_x formed in dry air flames. At temperatures around 2800°F the decrease is close to a factor of 8 while at higher temperatures it is around a factor of 3. These calculations indicate that NO_x formed decreases in flames with added moisture as compared to dry flames even when the T_{eq} for the two flames is the same. This suggests that all the reduction in NO_x cannot be attributed to thermal effects (as defined by T_{eq}).

The decrease in the ''Zeldovich and N_2O'' NO_x can be attributed to the reduction in O atom concentration. As mentioned earlier in the paper, at the same equivalence ratio the high humidity flames are of lower temperatures than dry flames. Hence, in order to evaluate the different flames at the same temperature, high humidity flames had to be at a higher equivalence ratio than dry flames. Scarcity of O_2, and, hence, lower O atom in the higher equivalence ratio humid flames would lead to lower NO_x formation by the ''Zeldovich and N_2O Mechanism.'' Humidity promotes $O+H_2O$ reaction, and, hence, suppresses O-atom concentration. A reduction in the ''Fenimore'' NO_x can be attributed to the change in the OH radical concentration in the flame. Higher OH concentration in humid air flames promotes oxidation of the hydrocarbon radicals [5], which would reduce ''Fenimore'' NO. Similar behavior was seen in the 5 percent side pilot flame too (Fig. 11). Due to the side pilot, NO_x levels were higher than the flame with no side pilot. The increase in NO_x for the flame with 5 percent fuel coming in from the side pilot was more predominant at the lower temperatures and was mostly due to higher Fenimore NO at the lean conditions (especially for the dry flame).

The factors mentioned in the introduction were also considered. There is definitely some reduction in NO_x due to lowering of temperature. Reduction of the peak temperature of the side pilot can also reduce NO_x formation, though the analysis with 0 percent side pilot indicates that this may not be the dominant factor. Suppression and delay of heat transfer rates due to added moisture would slow down the combustion process and lower super-equilibrium O-atom concentration and temperatures in reactors 2 and 3. These two reactors account for most NO_x formation and a lower temperature in these reactors (even though T_{eq} for the whole flame could be the same as the case with dry air) would lead to lower overall NO_x. Lower super-equilibrium O-atom concentrations would also lead to lower NO_x (reduction in ''Zeldovich and N_2O'' NO_x).

The percentage of NO reduction from the three mechanisms was similar for dry run and for humid air run. This indicates that even with moisture addition, which has the effect of lowering T_{eq}, the thermal NO channel remained the most dominant NO formation channel for most of the conditions examined in this study.

Conclusions

Data were collected over a wide range of equivalence ratio, moisture in the air stream, and pilot levels. The trends in the measured data have been used to evaluate the effect of various parameters on NO_x and CO emissions. The data have also been modeled using a PSR network and a literature reaction set. Good match between the data and modeling suggests that the approach used for modeling the combustor is useful. Addition of moisture to the air stream brings about a significant reduction in NO_x emissions. This reduction in NO_x is primarily due to a decrease in the T_{eq}, but reduction in NO_x with increasing moisture is also observed for flames in which T_{eq} is held constant. Analysis of the reaction set and computations for the three different NO_x formation pathways indicate the importance of O-atom concentration on NO_x formation. Moisture loading of 15 percent in the air reduces NO_x emissions by as much as a factor of 3 to 8 (depending on the T_{eq}) at the same T_{eq} as a dry flame. This reduction is attributed to the decrease in O-atom concentration. Even with 15 percent moisture content, most of the NO_x formation takes place via the ''Zeldovich'' pathway.

Acknowledgements

This work was supported by DOE/METC under contract DE-AC21-96MC33084. The authors would also like to thank Fred Robson, Brian Knight, and Archer Jennings for helpful discussions.

References

[1] Rao, A. D., 1989, ''Process for Producing Power,'' US Patent 4,829,763, May 1989.
[2] Humid Air Turbine (HAT) Cycle Public Report, Turbo Power and Marine (TPM), April, 1993.
[3] Robson, F. L., 1993, ''Advanced Turbine Systems Study-System Scoping And Feasibility Study,'' DOE Contract DE-AC21-92MC29247, United Technologies Research Center, East Hartford, CT.
[4] Dryer, F. L., 1976, ''Water Addition to Practical Combustion Systems-Concepts and Applications,'' *Sixteenth Symposium (International) on Combustion*, The Combustion Institute, Philadelphia, PA.
[5] Miyauchi, Y., Mori, Y., and Yamaguchi, T., 1981, ''Effect of Steam Addition on NO Formation,'' *Eighteenth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA.
[6] Touchton, G. L., 1985, ''Influence of Gas Turbine Combustor Design and Operating Parameters on Effectiveness of NO_x Suppression by Injected Steam or Water,'' ASME Paper 84-JPGC-GT-3.
[7] Blevins, L. G., and Roby, R. J., 1995, ''An Experimental Study of NO_x Reduction in Laminar Diffusion Flames by Addition of High Levels of Steam,'' ASME Paper 95-GT-327.
[8] Meyer, J.-L., and Grienche, G., 1997, ''An Experimental Study of Steam Injection in an Aeroderivative Gas Turbine,'' ASME Paper 97-GT-506.
[9] Kendrick, D. W., Anderson, T. J., Sowa, W. A., and Snyder, T. S., 1998, ''Acoustic Sensitivities of Lean-Premixed Fuel Injectors in A Single Nozzle Rig,'' ASME Paper 98-GT-382.
[10] Bowman, C. T., Hanson, R. K., Davidson, D. F., Gardiner, W. C., Jr., Lissianski, V., Smith, G. P., Golden, D. M., Frenklach, M., and Goldenberg, M., 1994, http://www.me.berkeley.edu/gri_mech/
[11] Glarborg, P., Kee, R. J., Grcar, J. F., and Miller, J. A., 1988, ''PSR: A Fortran

Compiler for Modeling Well-Stirred Reactors,'' Sandia National Laboratories Report, SAND86-8209, Livermore, CA.
[12] Leonard, G., and Stegmaier, J., 1994, ''Development of an Aeroderivative Gas Turbine Dry Low Emissions Combustion Systems,'' ASME Paper 93-GT-288.
[13] Maghon, H., Berenbrink, P., Termuehlen, H., and Gartner, G., 1990, ''Progress in NO_x and CO Emission Reduction of Gas Turbines,'' ASME Paper 90-IPGC/GT-4.
[14] Nguyen, Q. V., Edgar, B. L., Dibble, R. W., and Gulati, A., 1995, ''Experimental and Numerical Comparison of Extractive and In Situ Laser Measurements of Non-Equilibrium Carbon Monoxide in Lean-Premixed Gas Combustion,'' Combust. Flame, **100**, pp. 395–406.
[15] Reynolds, W. C., 1981, ''STANJAN,''Interactive Computer Programs for Chemical Equilibrium Analysis, Stanford University, Stanford, CA.
[16] Kee, R. J., and Lutz, A., private communication.
[17] Zeldovich, J., 1946, ''The Oxidation of Nitrogen Combustion and Explosions,'' Acta Physicochim. URSS, **21**, p. 577.
[18] Bowman, C. T., and Seery, D. J., 1972, *Emissions from Continuous Combustion Systems*, W. Cornelius and W. G. Agnew, eds., Plenum, New York, p. 123.
[19] Fenimore, C. P., 1971, ''Formation of Nitric Oxide in Premixed Hydrocarbon Flames,'' *Thirteenth Symposium (International) on Combustion*, The Combustion Institute, Philadelphia, PA.

Passive Control of Combustion Instability in Lean Premixed Combustors

Robert C. Steele

Luke H. Cowell

Solar Turbines Inc.,
San Diego, CA 92186-5376

Steven M. Cannon

Clifford E. Smith

CFD Research Corporation,
Huntsville, AL 35805

A Solar fuel injector that provides lean premixed combustion conditions has been studied in a combined experimental and numerical investigation. Lean premixed conditions can be accompanied by excessive combustion driven pressure oscillations which must be eliminated before the release of a final combustor design. In order to eliminate the pressure oscillations the location of fuel injection was parametrically evaluated to determine a stable configuration. It was observed that small axial changes in the position of the fuel spokes within the premix duct of the fuel injector had a significant positive effect on decoupling the excitation of the natural acoustic modes of the combustion system. In order to further understand the phenomenon, a time-accurate 2D CFD analysis was performed. 2D analysis was first calibrated using 3D steady-state CFD computations of the premixer in order to model the radial distribution of velocities in the premixer caused by non-uniform inlet conditions and swirling flow. 2D time-accurate calculations were then performed on the baseline configuration. The calculations captured the coupling of heat release with the combustor acoustics, which resulted in excessive pressure oscillations. When the axial location of the fuel injection was moved, the CFD analysis accurately captured the fuel time lag to the flame-front, and qualitatively matched the experimental findings. [S0742-4795(00)01103-0]

Introduction

Lean premixed combustion technology at Solar, referred to as SoLoNOx, began in the mid 1980s [1,2] with an initial guarantee of 42 ppmv (15 percent O2, dry) NO_x and a final design goal of 25 ppmv. The basic fuel injector concept developed was comprised of a swirler and parallel mixing duct into which natural gas is injected through fuel injection spokes. The fuel mixes with the swirling inlet air to produce a homogeneous gas and air mixture which is injected into the combustion chamber. The injector design includes a pilot fuel circuit which enables a portion of the fuel to be burned in the combustor in a diffusion-like flame. One of the original purposes of the pilot flame was to overcome combustion-driven pressure oscillations at low power operation.

Early development testing of the SoLoNOx systems revealed that the pressure oscillations were unacceptably high with the initial fuel injector designs. An increase in pilot fuel flow would suppress the pressure oscillations but result in an increase in NO_x emissions as shown in Fig. 1. In order to achieve the goal of 25 ppmv the design of the injector premixing section was modified to decouple the periodic release of energy with the acoustic modes in the combustion chamber. The well-known Rayleigh criterion [3] states that oscillations will be likely if the changes in heat release are in phase with the acoustic pressure disturbances. The recent works of Richards and Janus [4] and Lieuwen and Zinn [5] provide excellent background to an understanding of lean premixed combustion driven oscillations.

In order to eliminate the pressure oscillations associated with the Centaur family of engines [6], a modification to the annular premixing chamber was made as shown in Fig. 2. A comparison of the old and new designs reveals an added step in the outer wall which increases the residence time of the fuel and air mixture in the premix chamber and decouples the release of energy within the combustor. Without the need for high levels of pilot fuel to control the combustion oscillations, the NOx emissions were reduced to below 25 ppmv.

Similar results have been obtained for the Mars turbine fuel injector by adjusting the transport time of fuel and an mixture from the premix duct to the flame within the combustor. This was achieved by optimizing the axial location of fuel injection within the injector to decouple the release of energy with the combustion acoustics. The final design was determined by performing a series of experiments in a single injector facility at Solar. The engine pressure pulsations reported by Etheridge [7] during the initial field demonstration were eliminated without compromising the final emissions goal of 25 ppmv NO_x.

In addition to the use of experiments to control combustion driven oscillations, computational models are becoming more useful [8,9] in providing fundamental understanding of the physics associated with the phenomenon. These one-dimensional models are useful, but they cannot accurately assess the effects of geo-

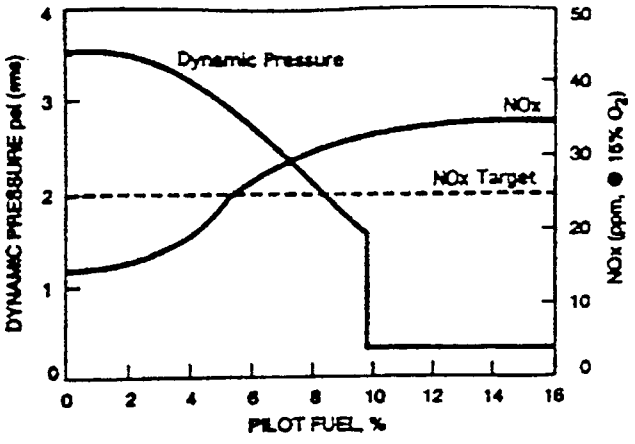


Fig. 1 Effects of pilot fuel on combustor pressure oscillations and NO_x emissions

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-52. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

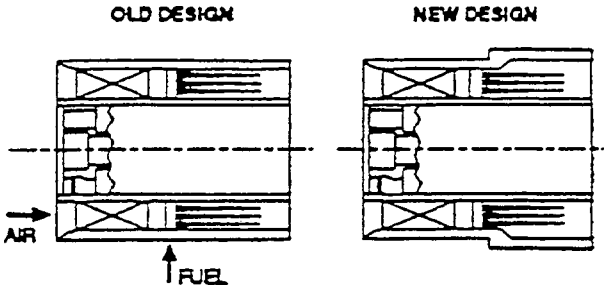


Fig. 2 Fuel injector modifications to eliminate combustor oscillations

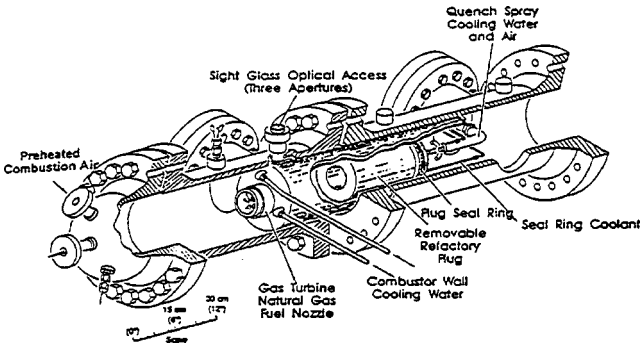


Fig. 4 Cutaway view of FETC test combustor [12]

metric details nor two-dimensional timelag effects caused by spatial variations in velocities and flame shapes. Two-dimensional computational models have been used to capture heat release and acoustic coupling in lean premixed combustors [10,11]. These studies investigated the effect of premixed fluid velocity and temperature on combustion instability. Good comparisons of predicted pressure oscillations to measurements obtained at the Federal Energy Technology Center were obtained.

In this paper a time-accurate two-dimensional CFD analysis is described and the numerical solutions are compared to a series of experiments conducted at Solar. The computational results are presented and model limitations discussed.

Experimental Setup and Procedure

The Mars premixer is used in Solar's low emissions, natural gas fired Mars 100S (10.5 MW) and 90S (9.4 MW) gas turbine engines [7]. A schematic of a basic research injector of this type is shown in Fig. 3. Here, natural gas is injected 1.17 premixing duct diameters upstream of the premixer exit. The fuel is injected into swirling air via injection spokes extending radially towards the premixer centerbody. The fuel and air mix in the portion of the premixing duct downstream of the fuel injection spokes, prior to injection into the combustion chamber. The injector also includes a pilot fuel circuit which enables a portion of the fuel to be burned in the combustor with a diffusion-like flame.

The Mars turbine premixer experiments were performed in the test facilities of Solar Turbines. In the test rig configuration, the combustor inlet air is preheated to a specified inlet air temperature prior to entering the test section. The combustor pressure and air flow-rate are controlled through modulation of valves upstream and downstream of the test section. The combustor is equipped with a refractory plug similar to the facility described by Richards et al. [12] and shown in Fig. 4. The combination of the combustion zone volume and the mass of gas in the exhaust neck approximates a classic acoustic Helmholtz resonator. The refractory plug is sized to match the critical combustor frequency experienced in the gas turbine engine. The natural frequency of the Mars combustor is approximately 350 Hz. The combustor dynamic pressure was recorded with a Kistler pressure transducer located outside the combustor rig. Combustor exhaust sampling was done using an area-averaging rake located downstream of the combustor exit plane. Overall combustor equivalence ratios were calculated from the O_2 measurements and agree to better than $+/-5$ percent of those calculated using the fuel and air flow-meter readings.

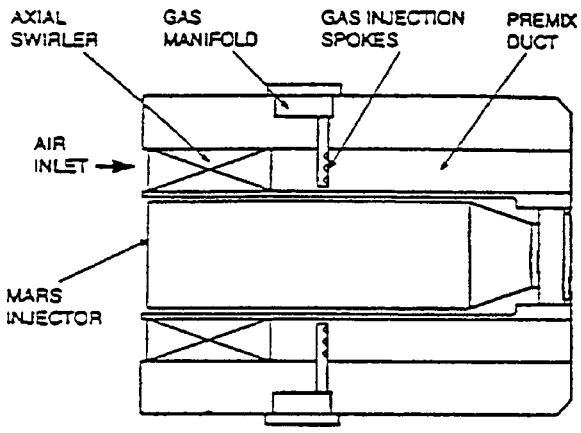


Fig. 3 Basic research injector design

Experimental Results and Discussions

A series of experiments, at a combustor pressure of 13.3 atm, an inlet temperature of 694 K (790°F), and a main injector equivalence ratio of 0.62, were conducted to determine the effect of fuel injection location on combustion instabilities. Mars turbine development fuel injectors were fabricated with fuel spokes located at different axial locations measured from the exit plane of the premix duct. The axial locations of the spokes were 3.5 cm (1.38 in), 4.1 cm (1.63 in), 4.5 cm (1.75 in), 5.7 cm (2.25 in), 5.8 cm (2.30 in), 6.9 cm (2.70 in), and 7.6 cm (3.0 in). The locations were chosen to substantiate the coupling of thermal energy with the acoustics of the combustor through the sinusoidal cycle of the pressure oscillation. It is explained by Putnum [13] that there is a time lag between the pressure, $P(t)$, and the heat release, $Q(t)$, which will result in a stable or unstable combustion system. The time lag is estimated as the distance between the point of fuel injection and flame front divided by the average axial velocity in the premix duct. In practice, it is not necessary that the heat release and pressure be exactly in phase to drive oscillations. Some driving will occur for heat release that leads or lags the pressure fluctuations by as much as 1/4 of the acoustic cycle [4]. There exists the opportunity to lead or lag the heat release such that there is combustion stability for ''short'' axial locations and ''long'' axial locations of fuel injection. In addition, the combustion system will remain stable until the heat release has shifted a complete acoustic cycle as described by Lieuwen and Zinn [5].

Presented in Table 1 are the test data for seven Mars turbine fuel injectors with unique spoke locations. The time lag (τ) is calculated assuming a bulk duct velocity of 45.7 m/s (150 ft/s) and a length of only the premixing section. The frequencies (F) are measured within the combustor.

The data indicate a shorter lag time results in a higher coupling frequency and a smaller τ^*F. Similarly, a longer lag time results in a lower coupling frequency and a larger τ^*F. At a τ^*F less that 0.30 and greater than 0.48 there exist regions of stability which have been confirmed on production Mars 100S and Mars 90S gas turbine engines. Shown in Fig. 5 are the regions of stability and instability for the Mars fuel injector. The lag time domain of thermal energy coupling is from one stable region to the other region or 0.91 ms. Given that the natural frequency of the combustor is 350 Hz and one cycle is 2.86 ms, the coupling of

Table 1 Comparison of pressure oscillations for different axial locations

Axial Loc. (cm)	Axial Loc. (in)	Time Lag (τ) (s)	Frequency (F) (Hz)	Pressure Oscillation	τ*F
3.5	1.38	0.76×10^{-3}	390**	No	0.30
4.1	1.63	0.90×10^{-3}	370	Yes	0.33
4.5	1.75	0.97×10^{-3}	360	Yes	0.35
5.7	2.25	1.25×10^{-3}	333	Yes	0.42
5.8	2.30	1.28×10^{-3}	313	Yes	0.40
6.9	2.70	1.50×10^{-3}	315	Yes	0.47
7.6	3.00	1.67×10^{-3}	290**	No	0.48

(** Low level pressure amplitudes indicating the limiting coupling frequency)

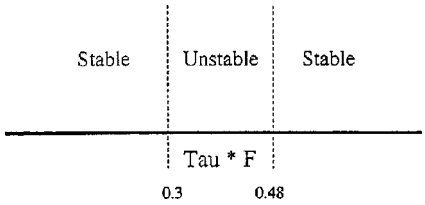


Fig. 5 The stable and unstable regions of τ^*F taken from injector tests

energy makes up 32 percent of the cycle. The experimentally determined coupling of +/−16 percent of the cycle falls short of the potential +/−25 percent of the cycle [13] due to acoustic damping of the overall combustion system. The effectiveness of the lag-lead time design approach is dependent on the stability of the flame as it is anchored to the fuel injector centerbody. The travel time of the fuel-air mixture from the exit of the premix duct to the flame front is assumed constant. The assumption allows the lag or lead time of the fuel and air mixture to be a function of the axial location of the fuel spokes.

CFD Model Definition

Time-accurate CFD modeling was utilized to further understand the driving mechanism for combustion instability in lean premixed combustors. Also, a comparison of predictions with instability measurements was utilized to help determine if the model could be used as a tool in the design of stable-operating combustor systems.

The interaction of acoustics, fluid flow, heat transfer, and chemical reaction was modeled by solving conservation equations for mass, momentum, energy, and chemical species using the CFD−ACE$^+$ software. The calculation domain started at the inlet plenum with a fixed mass flow inlet, and extended to the exit plane of the exhaust duct (fixed pressure), as shown in Fig. 6. In the actual experiments a fixed mass inlet does occur at the choked feed to the plenum and choked flow exists at the exhaust duct valve. The details of the choked feed to the plenum and the exhaust duct exit were not modeled, but the correct volumes and lengths for the inlet plenum/exhaust duct were utilized in the model. Because a Helmholtz (bulk) acoustic instability was excited, these boundaries have a negligible input on the instability prediction. The main geometrical features of the premix passage, combustor, and neck were obtained from the combustor hardware described in Figs. 3 and 4.

The swirl vanes were modeled as source terms that modified the axial momentum, swirl velocity, turbulent kinetic energy, and dissipation equations. A user subroutine applied the source terms to the computational cells at the swirler discharge. The change in total pressure across the swirl vanes was assumed to be zero. The inclusion of these swirl source terms within the computational domain allowed acoustic pressure waves to effectively travel through the swirl vane passages.

The walls were specified as adiabatic, no-slip boundaries. Methane fuel was injected through three rings at radial locations simulating the fuel orifice locations. The fuel was injected at an axial station upstream of the combustor dump plane. A premixed methane-air mixture was also injected through a pilot orifice at the center of the combustor dump plane. Cooling air was injected along the combustor dome and liner walls. Fixed mass flow rates were specified at all inlets except the pilot, where a total pressure was specified. The grid was heavily concentrated in the premix passage and the first section of the combustor to accurately capture the fuel injection and flame front. Many-to-one grid topology was used at the inlet to the premix passage and between the two sections of the combustor. The total grid was 11,261 computational cells. This was equivalent to 20,468 computational cells for a non-many-to-one grid.

The CFD−ACE$^+$ code was used to perform the calculations with the following numerics and physical models:

1 second order Crank-Nicholson scheme for temporal differencing;

2 a smart scheme for spatial differencing of the following variables: u, v, and w velocities, pressure correction, turbulent kinetic energy (k) and dissipation (ε), enthalpy, mixture fraction, and progress variable. The smart scheme automatically chooses between first order upwind, second order upwind, and central differencing depending on local gradients;

3 renormalization group (RNG) k-ε turbulence model [14] and standard wall functions. The RNG k-ε model has been shown to give reasonable engineering time-accurate solutions, (e.g., see [15]) as long as the turbulent time scale k/ε is less than the time scale of the large scale motion; and

4 one-step reaction rate equation (methane reacting with air to equilibrium products). By going to equilibrium products, a more accurate flame temperature and overall heat release is predicted than if the products are restricted to CO_2 and H_2O. The reaction rates had a form similar to Westbrook and Dryer [16], but were calibrated in CFD−ACE$^+$ to laminar flamespeeds and autoigni-

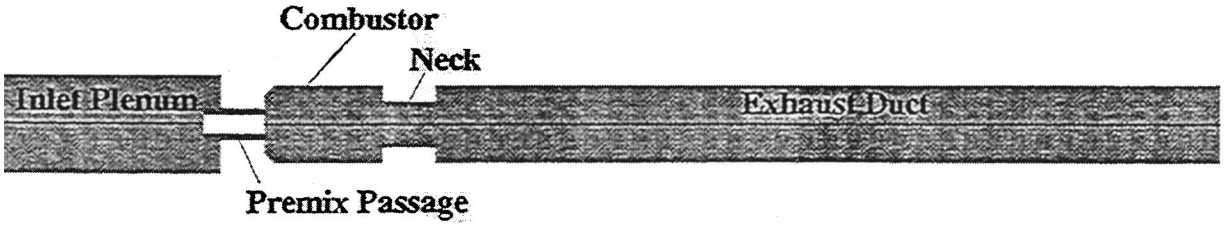


Fig. 6 Calculation domain for CFD analysis

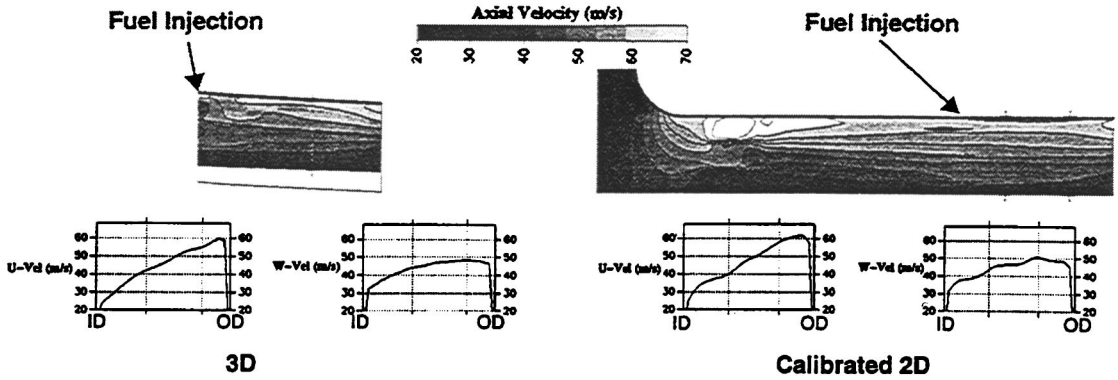


Fig. 7 Comparison of predicted velocity distributions in the three-dimensional and calibrated two-dimensional premixer models

tion delay times of methane-air combustion. Steady-state calculations were also performed using the 5-step global mechanism of Malte and Nicol [17] for high pressure, lean premixed methane combustion. The five-step global mechanism accounts for the super-equilibrium formation and subsequent oxidation of CO as well as the formation of NO_x due to thermal, prompt, and nitrous oxide pathways.

Calibration of Two-Dimensional Model. Three-dimensional steady-state CFD calculations of the premixer were used to calibrate the radial distribution of velocity in the two-dimensional model. This calibration was important due to the strong effect of premixer velocity distributions on flame location and timelag. A full three-dimensional CFD model can capture the effect of flow restriction just upstream of the swirl vanes and swirling flow aspects.

A periodic 18 degree sector of the full three-dimensional annular passage was modeled. The three-dimensional model included the physical representation of the airflow dump into premixer, swirl vane and fuel spoke. The fuel dump at the entrance to the premix passage was originally modeled as a backstep in the two-dimensional calculation. The recirculation zone downstream of the backstep extended to the swirl vanes. This caused unrealistic backflow through the vanes. Unlike the two-dimensional, axisymmetric predictions, the three-dimensional results showed a short recirculation zone length and positive flow at all radial locations along the swirler exit plane. To model this effect on the axial velocity, the backstep was eliminated from the two-dimensional domain and was instead replaced with a momentum resistance concentrated at the backstep location. The momentum resistance is calculated as

$$\Delta p = \frac{K_q \rho u^2}{2\beta^2},$$

where K_q is the quadratic resistance coefficient, ρ is the density, u is the velocity through the unrestricted area, and β is the porosity of the resistance (flow area/total area). A K_q value of 1 and β values of 0.15 and 0.7 over the lower and upper half of the backstep region respectively were used in the two-dimensional predictions.

In order to capture the correct tangential velocity effect of the swirl vane in the two-dimensional calculations, the swirl angle in the swirl source term was assumed to vary linearly between 69 deg at the inner diameter and 32 deg at the outer diameter. The effect of the swirler on turbulent mixing was also included in the two-dimensional model as a turbulence intensity of 10 percent and a length scale of 0.02 m (0.79 in) were specified at the swirler discharge. Figure 7 shows the predicted axial and tangential velocity profiles in the premixer using the three-dimensional and calibrated two-dimensional model. Good agreement in the velocity profiles was obtained downstream of the fuel injection spoke.

Steady-State CFD Calculations. Steady-state calculations were performed with the one-step CH_4 oxidation mechanism to equilibrium products and the five-step global reaction models. Figure 8 shows a comparison of heat release contours for the two cases. A fluid trace of 2.0 msec that starts 5.6 cm (2.22 in) upstream of the combustor dump plane is superimposed on the predictions. The predictions show the flame is anchored on the centerbody, and the flame is angled to the combustor centerline. The tuned one-step model accurately represents the heat release location. Only minor differences can be observed. Due to the faster computational times and the reasonable representation of the steady-state flame location, the one-step model was used for the subsequent time-accurate calculations.

Steady-state CFD results showed the swirling flow from the premix passage attached to the combustor wall and an outer and inner recirculation zone are formed. The Helmholtz resonant frequency for this combustor/neck geometry is 365 Hz, assuming a bulk gas temperature of 1600 K. The mass-averaged velocity of fluid in the premix passage is 45.7 m/s (150 ft/s), though the axial velocity varies from 30 to 60 m/s at the inner and outer diameters respectively. Figure 9 shows fluid traces in the steady-state flowfield starting at various fuel injection locations upstream of the combustor dump plane. The fluid traces end at a time lag corresponding to the Helmholtz characteristic time minus the acoustic time from the flame zone to the fuel injector location $(\tau_H - \tau_a)$. If this timelag ends at the heat release zone, then unstable combustion would be likely. The results show that, according to steady-state analysis, the fluid traces for fuel injection locations shorter than 3.8 cm (1.5 in) extend beyond the heat release zone. The case at 7.4 cm (2.9 in) shows most of the premixed fluid to be in phase with the steady-state heat release zone. The case at 9.5 cm (3.75 in) shows most of the premixed reactant fluid to be out of phase

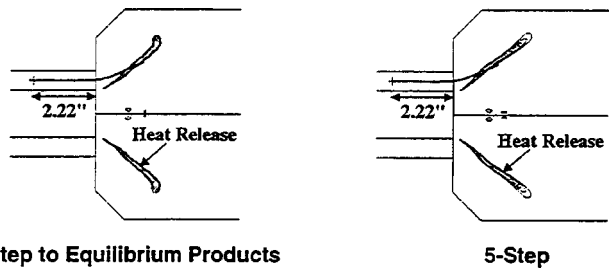


Fig. 8 Comparison of heat release using the one-step to equilibrium products and the five-step chemical reaction models

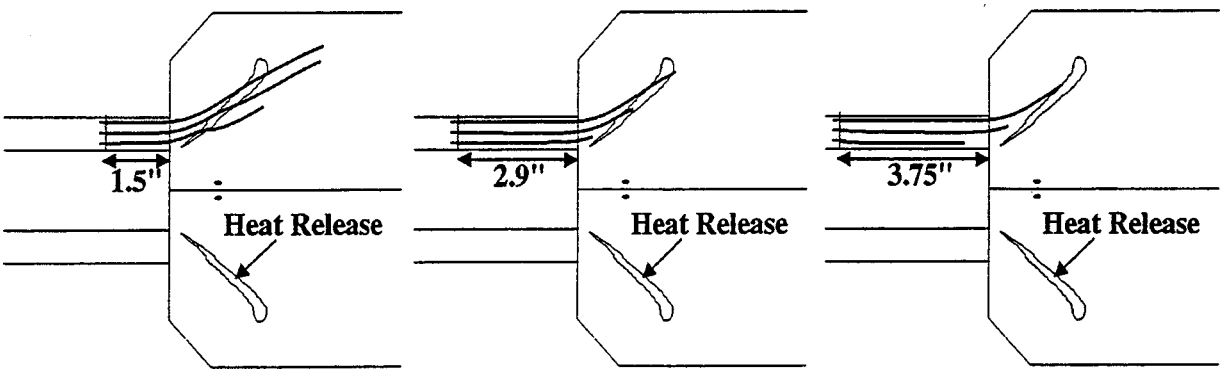


Fig. 9 Predicted mean fluid traces for different fuel injection locations

with the steady-state heat release zone. These results also indicate that the timelag from the fuel injection location to the flame zone is different for inner diameter reactants and outer diameter reactants. This two-dimensional effect on timelag is difficult to capture in one-dimensional instability models, since a priori knowledge of the flame location and velocity distribution is required. Of course, the timelag will be different once the flame begins to move during unsteady combustion. Time-accurate CFD calculations can capture the initial time-lag driven instability caused by starting from a steady-state as well as the shift in timelag during the onset of instability and the ensuing flame motion.

Time-Accurate CFD Calculations. The time-accurate calculations were performed at a time step of 1.6e-5 s and typically required between 12 and 24 CPU hours on a DEC Alpha 500 workstation. Starting from converged steady-state calculations, time-accurate CFD analyses were performed for different fuel injection locations. The initial pressure disturbance caused by starting from a steady-state solution can grow into a large amplitude pressure oscillation if the unsteady heat release couples with natural resonant acoustics, or it can decay to a small value if no coupling exists. The time-accurate CFD model predicts the evolution (including amplitude and frequency) of pressure, velocity, temperature, enthalpy, mixture fraction, and progress variable. Time-accurate calculations were performed for fuel injection locations at 4.3 cm (1.69 in), 4.6 cm (1.81 in), 4.9 cm (1.93 in), 5.8 cm (2.29 in), 6.7 cm (2.65 in), 7.6 cm (3.0 in), 7.9 cm (3.12 in), and 8.5 cm (3.36 in) upstream of the combustor dump-plane.

To demonstrate the predicted unstable combustor behavior, the pressure history corresponding to one set of unstable conditions

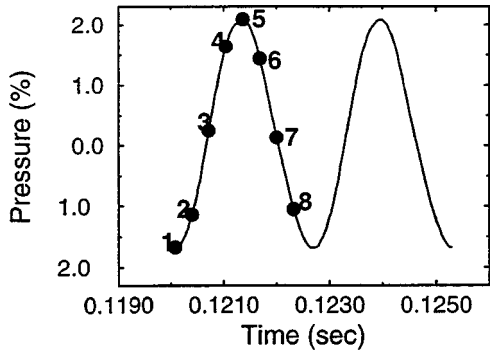


Fig. 10 Predicted limit cycle pressure history at combustor mid-section [fuel location=5.8 cm (2.29 in)]

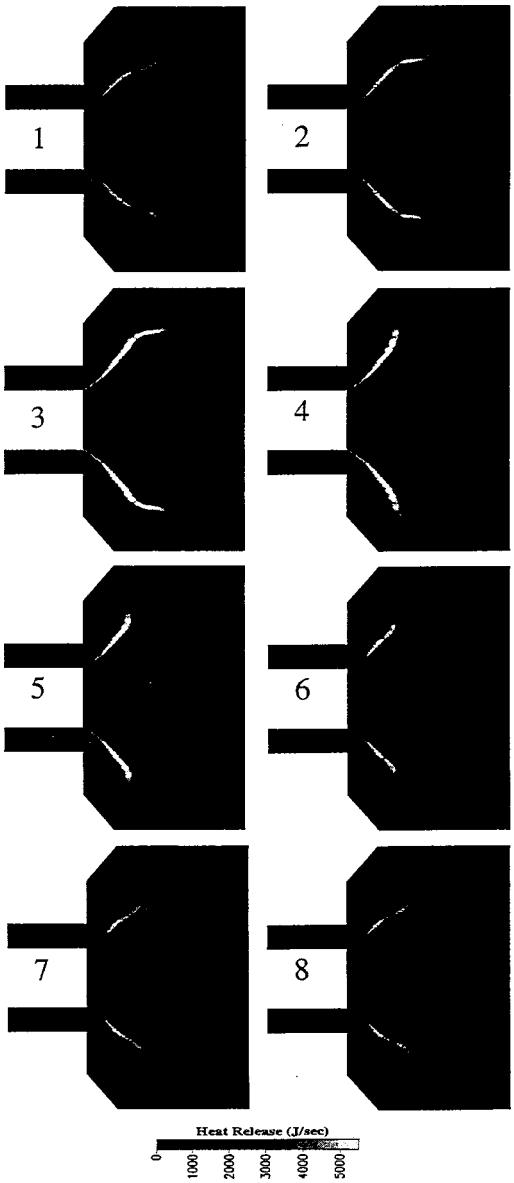


Fig. 11 Predicted heat release contours during a complete cycle [fuel location=5.8 cm (2.29 in)]

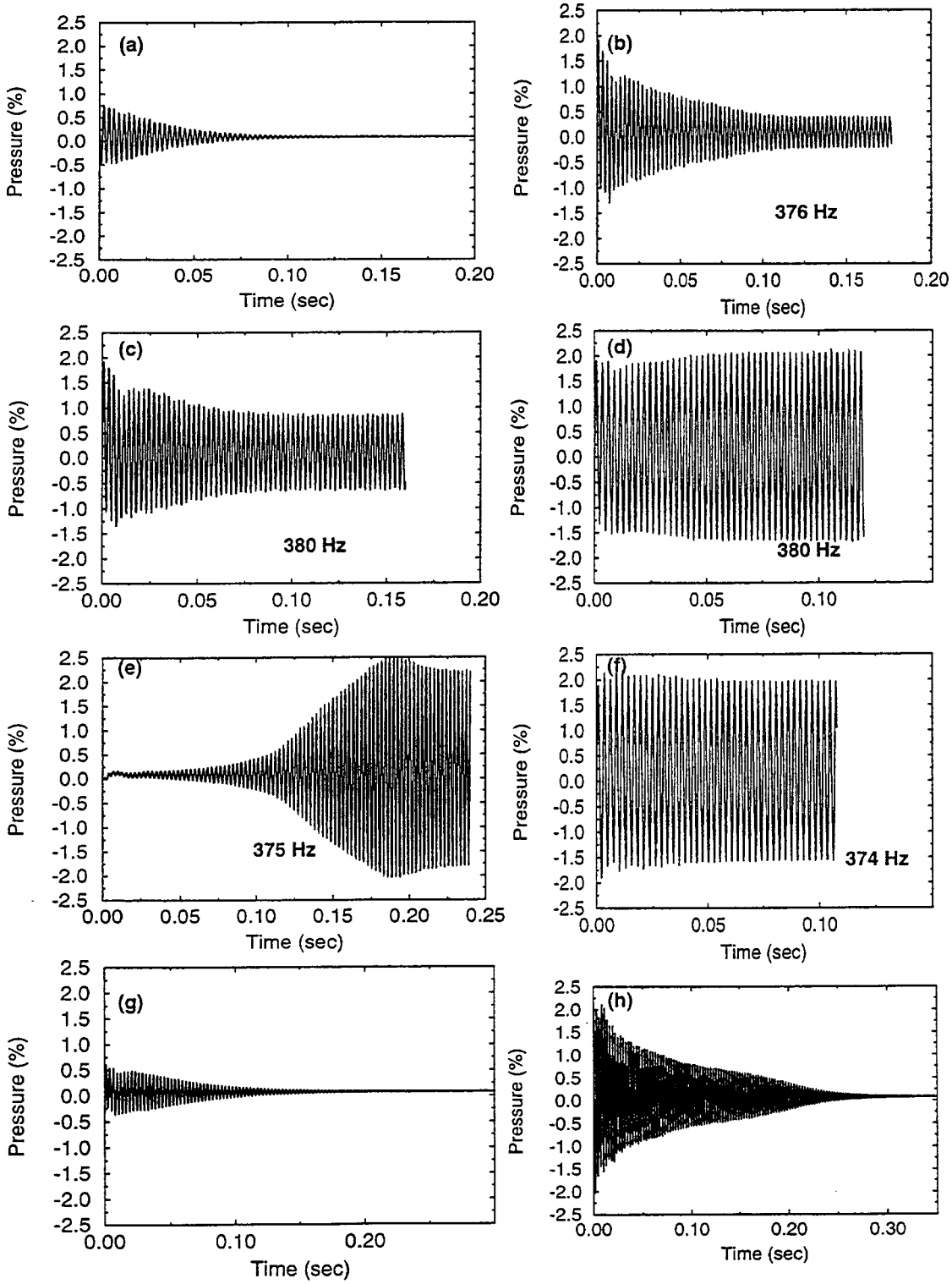


Fig. 12 Time history of pressure at combustor mid-section for fuel injection locations of (*a*) 4.3 cm [1.69 in], (*b*) 4.6 cm [1.81 in], (*c*) 4.9 cm [1.93 in], (*d*) 5.8 cm [2.29 in], (*e*) 6.7 cm [2.65 in], (*f*) 7.6 cm [3.0 in], (*g*) 7.9 cm [3.12 in], and (*h*) 8.5 cm [3.36 in]

(5.8 cm (2.29 in)) is shown in Fig. 10. Also shown are eight equally spaced points (0.34 msec) in a complete cycle. The predicted pressure oscillates within the combustor at a frequency of 380 Hz. The theoretical Helmholtz frequency is 376 Hz (assuming the average predicted combustor temperature of 1670 K) and the measured frequency varied between 370 and 315 Hz. This good agreement in frequency indicates that the model is adequately capturing the temperature distribution within the combustor and the predicted dynamic boundary conditions for the Helmholtz geometry. The existence of large scale oscillations also indicate that the model is capturing the driving effect of unsteady heat release for this case.

This unstable behavior causes the flame to exhibit a radial and axial back-and-forth motion. Figure 11 shows predicted heat re-

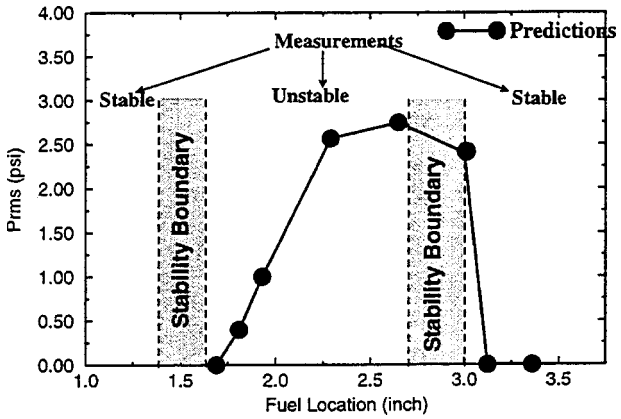


Fig. 13 Comparison of measured and predicted instability results as a function of fuel injection location

lease contours in the combustor for the eight instantaneous moments (Fig. 10) during the limit cycle. Most of the heat release occurs at the inner shear layer, between the premixer reactants and inner recirculation zone products. At points 1 and 2 in the cycle, the pressure is a minimum and the axial velocity of reactants into the combustor is increased. The maximum axial penetration of the heat release zone occurs at point 3, where the pressure begins to reach a peak. At the time of peak pressure (4 and 5), the flow of reactants into the combustor is decreased, but the heat release zone has moved out radially and the local heat release reaches a maximum. As the pressure decreases again (points 7 and 8), the maximum heat release zone moves back towards the combustor centerline and higher velocity reactants are allowed to enter the combustor.

As explained by Janus et al. [18], the advection time for fluid travel down the nozzle duct, the mixing time between reactants and products, and the chemical reaction time can be combined to obtain an overall transport time. As can be seen, the advection time is dominant in this calculation. If this transport time matches or is near an acoustic resonance and if the acoustic gain is greater than the acoustic loss, then combustion instability will result. Changes in the fuel injection location had a large effect on the unsteady combustor behavior. Figure 12 shows the time history and frequency of the pressure at the combustor mid-length for the cases with different fuel injection locations. The 5.8 cm (2.29 in) case was started from a steady-state solution, while the remaining cases were started from the transient 5.8 cm (2.29 in) case. A change in the fuel injection location from 4.3 to 5.8 cm changed the mean advection time by 0.33 msec ($\sim$1/8 of timelag). Thus the transport time was changed enough to be in phase with the Helmholtz resonant frequency and large-amplitude pressure oscillations resulted. When the fuel injection location was moved from 5.8 to 7.9 cm, the mean advection time changed by 0.46 msec (between 1/6 and 1/7 of timelag), enough to cause the unsteady heat release and acoustics to be out of phase, and stable combustion was again obtained. The calculated frequencies varied only slightly between 374 and 380 Hz for the different fuel injection locations.

This predicted combustor behavior as a function of fuel injection location was similar to that observed in the experimental tests. Figure 13 shows a comparison between predicted and measured rms pressure results as a function of fuel injection location. The absolute range of unstable fuel injection locations was identical to the measurements. The instability region was shifted upstream by only 0.12 in. compared to the measurements. If the heat release region were shifted downstream, then better agreement may be expected.

The CFD model was not able to mimic the variation in frequency with fuel injection location as shown experimentally. The predicted instability was always near the same frequency ($\sim$375 Hz), regardless of the fuel injection location. The measured frequency, on the other hand, varied between 315 and 370 Hz. It is not clear why the CFD results exhibit no change in frequency over the range of coupling. A more expensive simulation with a finer grid, more detailed modeling of subgrid turbulence effects, and higher order temporal and spatial differencing may be required to describe the variation in oscillating combustor frequency.

Conclusions

An experimental and numerical investigation of passive combustion instability control in an industrial dry low-NO_x combustor was presented. Measurements indicated a strong effect of fuel time lag on combustion instability. Large amplitude pressure oscillations were avoided by moving the fuel injection to certain axial locations within the premixer.

Two-dimensional time-accurate CFD analysis of the lean premixed combustion system was used to further understand the effect of fuel time lag on combustion instability. The velocity distribution in the premixer was affected by non-uniform inlet flow conditions and was captured in a steady-state three-dimensional CFD analysis. The three-dimensional analysis was used to calibrate the premixer velocity distribution for a two-dimensional CFD model. Subsequent two-dimensional time-accurate CFD calculations captured the coupling of unsteady heat release with Helmholtz acoustics. When the axial location of fuel injection was moved, the model captured the fuel time lag to the flamefront, in agreement with the experimental data. The two-dimensional time-accurate CFD analysis can be applied to the design of real engines if the following are provided: (1) the frequency of oscillation for the real combustor and (2) the premixer velocity profiles from a detailed three-dimensional model. This allows the two-dimensional time-accurate model to describe the real engine acoustics (with a Helmholtz geometry) and fuel time-lag that may drive combustor oscillations.

The qualitatively accurate results and the low computational time show the promise of using this numerical approach to assess passive control strategies (geometric changes) for lean premixed combustors. The assessment of active control strategies is also possible.

Acknowledgments

The authors wish to acknowledge the support of the manufacturing department and the development test group at Solar for their support of the project. We are grateful for the helpful discussion with Dr. David Black of CFDRC and for the effort of Marni Kent at CFDRC in accurately preparing this typescript.

References

[1] Roberts, P. B., Kubasco, A. J., and Sekas, N. T., 1981, ''Development of a Low NOx Lean Premixed Annular Combustor,'' ASME paper 81-GT-40.
[2] Smith, K. O., Angello, L. C., and Kurzynske, F. R., 1986, ''Design and Testing of an Ultra-Low NOx Gas Turbine Combustor,'' ASME paper 86-GT-263.
[3] Rayleigh, J. S., 1878, ''The Explanation of Certain Acoustical Phenomena,'' Royal Institute Proceedings, **VIII**, pp. 536–542.
[4] Richards, G. A., and Janus, M. C., 1997, ''Characterization of Oscillations During Premix Gas Turbine Combustion,'' ASME Paper 97-GT-244.
[5] Lieuwen, T., and Zinn, B. T., 1998, ''The Role of Equivalence Ratio Oscillations in Driving Combustion Instabilities in Low NOx Gas Turbines,'' presented at the 27th International Symposium on Combustion, Boulder, CO, August.
[6] Rawlins, D. C., 1995, ''Dry Low Emissions: Improvements to the SoLoNOx Combustion System,'' Eleventh Symposium on Industrial Applications of Gas Turbines Canadian Gas Association, Banff, Alberta, Canada, October.
[7] Etheridge, C. J., 1994, ''Mars SoLoNOx—Lean Premix Combustion Technology in Production,'' ASME Paper 94-GT-255, The Hague.
[8] Paschereit, C. O., and Polifke, W., 1998, ''Investigation of the Thermoacoustic Characteristics of a Lean Premixed Gas Turbine Burner,'' ASME Paper 98-GT-582.
[9] Peracchio, A. A., and Proscia, W. M., 1998, ''Nonlinear Heat-Release Acoustic Model for Thermoacoustic Instability in Lean Premixed Combustors,'' ASME Paper 98-GT-269.
[10] Smith, C. E., and Leonard, A. D., 1997, ''CFD Modeling of Combustion

Instability in Premixed Symmetric Combustors,'' ASME Paper 97-GT-305.
[11] Cannon, S. M., and Smith, C. E., 1998, ''Numerical Modeling of Combustion Dynamics in a Lean Premixed Combustor,'' FACT-Vol. 22, 1998 International Joint Power Generation Conference.
[12] Richards, G. A., Gemmen, R. S., and Yip, M. J., 1997, ''A Test Device for Premixed Gas Turbine Combustion Oscillations,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 776.
[13] Putnam, A. A., 1971, *Combustion Driven Oscillations in Industry*, American Elsevier Publishers, New York, NY.
[14] Yahkot, V., Orszag, S. A., Thangam, S., Gatski, T. B., and Speziale, C. G., 1992, ''Development of Turbulent Models for Shear Flows by a Double-Expansion Technique,'' Phys. Fluids, **4**, pp. 1510–1520.
[15] Spring, S. A., Smith, C. E., and Leonard, A. D., 1995, ''Advanced Demonstration of Fuel Injector/Flameholder for High Speed Ramburners,'' WL-TR-95-2068, CFDRC Report No. 4166/2.
[16] Westbrook, C. K., and Dryer, F. L., 1981, ''Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames,'' Combust. Sci. Technol., **27**, pp. 31–43.
[17] Malte, P. C., and Nicol, D. G., 1996, ''Development of a Five-Step Global Mechanism for Methane Oxidation and NO Formation,'' final report to Applied Science and Engineering Technologies and Solar Turbines, University of Washington, Seattle, WA, June 12, 1996.
[18] Janus, M. C., Richards, G. A., Yip, M. J., and Robey, E. H., 1997, ''Effects of Ambient Conditions and Fuel Composition on Combustion Stability,'' ASME Paper 97-GT-266.

Combustion Instabilities in Industrial Gas Turbines—Measurements on Operating Plant and Thermoacoustic Modeling

David E. Hobson
e-mail: dave.hobson@powertech.co.uk

John E. Fackrell

Gary Hewitt

PowerGen plc,
Power Technology Centre,
Ratcliffe-on-Soar,
Nottingham, NG11 0EE, UK

Measurements of vibration and combustion chamber dynamic pressures have been taken on a number of 150 MW industrial gas turbines operating on pre-mixed natural gas, both during long periods of base-load operation and during short duration load-swings. The data has been analyzed in terms of the frequency and bandwidth of the principle peak in the vibration and pressure spectra as a function of load and other operating parameters. It is observed that bandwidth, which is a measure of the damping of the resonant mode of the combustion chamber's acoustic resonance, decreases towards zero as the machines approach their combustion stability limits. A theoretical model of the thermoacoustic behavior of the combustion system has been developed to see to what extent the observed behavior on the operational machines can be explained in terms of an acoustic model of the ductwork and a flame characterized simply by a time-delay. This time delay is obtained from the frequency response function of the flame in response to unsteady perturbations in inlet velocity and is calculated using computational fluid dynamics. The model has also been used to illustrate the importance of fuel supply system design in controlling combustion stability. It is shown that stability can be a strong function of the acoustic impedance of the fuel supply and that this can lead to enhanced or reduced stability depending on the flame characteristics. [S0742-4795(00)01403-4]

1 Introduction

Many large industrial gas turbines operating on pre-mixed natural gas suffer from high levels of dynamic pressure or ''humming'' in the combustion chambers that can sometimes prevent them from reaching full load and can lead to serious damage to internal components. The high pressure levels are due to acoustic resonances inside the combustion chamber cavities which are coupled with perturbations in heat release at the flame. To help understand the phenomenon and in particular to identify the geometric and operating parameters on which the dynamic pressures depend, simple thermoacoustic models have been developed for the combustion chamber components including the fuel supply. In support of this, a CFD model of the swirling flow from a burner has been used to quantify the flame parameters affecting overall combustion stability, in particular the time lags between pre-mixed fuel injection and heat release in the flame. The model has been used to establish broad generic characteristics of combustion instabilities. In addition, it has been applied to a Siemens V94.2 gas turbine and used to quantify two characteristics of the humming signals, namely frequency and damping and their variation with load. The results have then been compared with data obtained on three operational gas turbines.

2 Modeling of Combustion Instability

A number of approaches have been taken to the modeling of thermoacoustic instability. Richards et al. [1] used a lumped parameter approach incorporating a bimolecular reaction rate law to produce a non-linear model of a pulsed combustor. This model was later adapted to look at lean pre-mixed combustion [2]. Smith and Leonard [3] applied an axisymmetric, transient CFD analysis of the full Navier Stokes equations to the same pre-mixed system. Transient CFD analyses of a suddenly perturbed flame have also been used to predict a so-called flame frequency response function (FFRF) for both a diffusion flame system, Bohn et al. [4], and a pre-mixed system, Bohn et al. [5]. A simpler quasi-one-dimensional CFD analysis was used by Liu and McGuirk [6] to predict the unsteady behavior of an after-burner model.

However, the most frequent approach to the problem has been through the use of the one-dimensional acoustic wave equations based on a linearized form of the conservation equations for mass, momentum, and energy. A good description of this approach has been given by Dowling [7] and a model based on this approach has been adopted for the current work. If the acoustic properties of the upstream and downstream parts of the combustion system are known, then the acoustic equations may be solved by any standard technique to give the fluctuating pressures and velocities either side of the flame front. In the present case, a transfer matrix technique is used, in the same manner as described by Chadha et al. [8] for a valve-pipework system. The acoustic pressures are assumed continuous across the flame front and the acoustic velocities can be related by the jump in volumetric flux associated with the heat release. Specification of an FFRF relating the fluctuating heat release to the fluctuating velocity closes the set of equations and allows a complete solution to be obtained.

The FFRF would normally have to be found from experiment, either a physical experiment or a numerical experiment such as that by Bohn et al. referred to above. However, many of the experimental results obtained (e.g., [9–11,4]) have shown that the form of the function can be characterized (at least approximately) by simply assuming a time delay or phase lag between the heat release and velocity fluctuations. This time delay function has been used in much, probably most, of the modeling of thermoacoustic instability that has been done (e.g. [11–13]) and its use is explored in the present work. The total delay time is made up of a number of contributions from the chemical kinetics and the fluid

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-110. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

dynamic convection and mixing, but in highly turbulent gas turbine combustors, it is the fluid dynamic convection time between fuel inlet and flame front which tends to dominate.

3 Thermoacoustic Model

In the Siemens V94.2, combustion takes place within two large silos, symmetrically placed either side of the machine centerline (see Fig. 1 taken from [14]).

At the compressor discharge, the air divides and the flow in each silo passes upwards along an annulus surrounding the flame tube. It turns and passes through eight burners fitted with swirl vanes and into the combustion chamber which is lined with ceramic tiles, and then downwards to the turbine entry. In premix mode, fuel is supplied through a series of distribution pipes immediately upstream of the swirl vanes. A pilot diffusion flame immediately downstream of the vanes is used to provide stability over a wide range of excess air ratios [14]. Pressure losses between compressor and turbine are very small and confined mainly to the swirl vanes. Small cooling air holes allow air to flow from the cold duct to the hot duct; these have not been incorporated into the full model but an estimate of their influence has been made.

3.1 Linearized Acoustic Equations. A simple thermoacoustic representation (Fig. 2) of the actual combustion system has been used in this work. It represents the simplest possible model for studying the V94.2 as well as for examining more generic aspects of combustion instabilities. To do this, a number of assumptions have been made as follows:

Small Perturbation Analysis. Only conditions prior to instability are considered; dynamic pressures are much less than 100 mbar and acoustic velocities less than 2 m/s. A linear, small perturbation analysis is therefore appropriate with, for example, pressure varying as

$$p(x,t) = P_c + p'(x) \cdot \exp(rt). \tag{1}$$

The variable r is complex with $\mathrm{Im}(r)=\omega$, the frequency of oscillation, and $\mathrm{Re}(r)=s$, the factor determining growth or decay of the oscillations. Generally, both the modeling and experiments show that $s \ll r$ for all cases of interest.

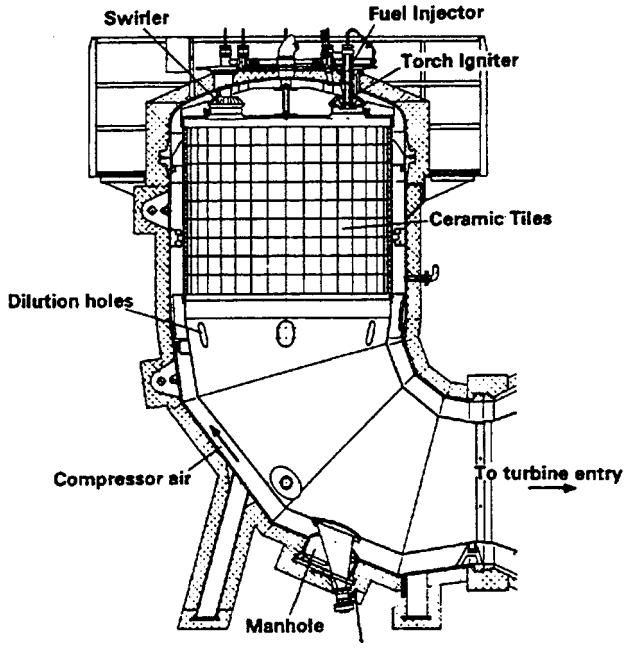


Fig. 1 Siemens V94.2 silo combustion chamber

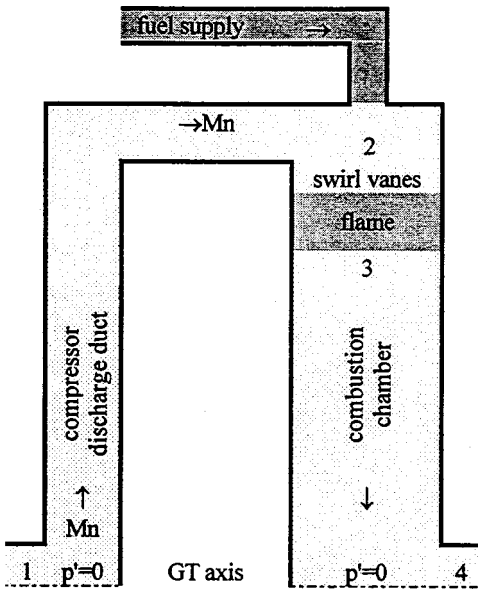


Fig. 2 Components of the acoustic model of the V94.2 combustion chamber silo

Boundary Conditions. It has been observed that at the main humming frequency, 75 to 78 Hz, pressures in the two silos are in anti-phase, implying that acoustic boundary conditions of zero unsteady pressure $p_1' = p_4' = 0$ can be set on the machine centerplane where the inlet and outlet flows from the silos come together. As a consequence of this, the acoustic impedances presented to the flow by the turbine inlet and compressor discharge are actually unimportant and can be ignored. Note though that these impedances will be important in controling other modes where p_1' and p_4' are not zero, though their very high values will tend to suppress acoustic velocity fluctuations into and out of the compressor and turbine.

Compressor Discharge Ductwork. This annulus is represented by a duct of equivalent length and (constant) area to the actual passages between the compressor discharge and the burner inlets. The flow of air in the duct is frictionless and at low Mach number and the acoustic behavior is assumed to be one-dimensional, though in general transverse waves cannot be ruled out. In the one-dimensional case, simple transfer matrix relationships can be derived relating the inlet and outlet unsteady flows and pressures:

$$\begin{vmatrix} p_2'/\rho c \\ u_2' \end{vmatrix} = \begin{vmatrix} \cosh(rL/c) & -\sinh(rL/c) \\ -\sinh(rL/c) & \cosh(rL/c) \end{vmatrix} \times \begin{vmatrix} p_1'/\rho c \\ u_1' \end{vmatrix}. \tag{2}$$

In this expression L, c, and ρ refer to the length of the inlet duct, the velocity of sound and density of the inlet air.

Combustion Chamber. Similarly, the hot zone is represented by a duct of length and area equivalent to the passages up to the turbine inlet guide vanes. A constant temperature is assumed throughout and the combustion products given the same properties as air. At this stage, cooling flows are ignored.

Swirl Vanes. All pressure losses (1.2 to 1.5 percent of absolute pressure) are assumed to be confined to the vanes (there is no pressure loss across the flame). The steady pressure drop is assumed to be equal to the dynamic head $^1/_2\rho_c U_s^2$ of the air leaving the vanes and the unsteady pressure loss to be quasi-steady, i.e., $\Delta p' = p_3' - p_2' = -\rho_c U_s u_s'$.

Fuel Supply. The fuel is assumed to be fully pre-mixed and with constant stoichiometry at all times.

Flame. For the fully pre-mixed fuel supply the unsteady heat release per unit inlet area Q' is directly proportional to the perturbation in unsteady velocity u_2' (see, e.g., [7]):

$$Q' = c_p(T_h - T_c)\rho_c u_2'. \tag{3}$$

The heat release results in an increase in volumetric flow rate given by

$$A_h u_3' - A_c u_2' = (\gamma - 1)/(\rho_c c_c^2)\cdot Q'. \tag{4}$$

Also, it is assumed that the distance from the swirl vanes to the flame is small compared to typical acoustic wavelengths and that there is no pressure loss across the flame.

Flame Frequency Response Function. Equation 3 is a quasi-steady approximation to the flame characteristics and is valid only at low frequencies. At higher frequencies, it must be modified by a frequency response function $\psi(\omega)$ so that

$$Q' = \psi(\omega)\cdot c_p(T_h - T_c)\rho_c u_2' \tag{5}$$

$\psi(\omega)$ tends to unity at low frequencies and can often be approximated by a simple time-delay function at higher values:

$$\psi(\omega) = \exp(-j\omega\Delta t). \tag{6}$$

The frequency response is the only potential source of instability so it is the time delay Δt or phase lag $\omega\Delta t$ between the unsteady velocity u_2' and the heat release Q' which is the most important ingredient in determining stability or otherwise. This will be considered in more detail later.

3.2 Solutions to Equations. The boundary conditions, the acoustic transfer matrices and flame equations can be gathered together into a 9 by 9 matrix $\mathbf{M}(r)$ relating the unsteady flows, pressures and heat release Q'.

$$\mathbf{M}(r)\times \begin{vmatrix} u_1' \\ p_1' \\ \cdot \\ \cdot \\ u_4' \\ p_4' \\ Q' \end{vmatrix} = |0| \tag{7}$$

This represents an eigenvalue problem in the unknown complex variable r, the solutions for which can be determined by setting the determinant of matrix $\mathbf{M}(r)$ to zero and solving for r. This has been accomplished using the symbolic mathematics program Mathcad for a wide range of temperatures and mode numbers up to $m=7$.

Examination of the imaginary part of r, $\mathrm{Im}(r)=\omega$, then determines the frequency of any thermoacoustic oscillations, stable or otherwise, whereas the ratio of the real part to the imaginary part gives a value for the damping ratio $\zeta = -\mathrm{Re}(r)/\mathrm{Im}(r)$. Positive values of ζ infer stability and vice versa. It is these parameters which will be compared with experimental values.

Finally, for any particular solution for r and an arbitrary value for u_1', it is possible to determine values of p_2', u_2', etc. and intermediate values of pressure and velocity, to plot the mode shape associated with any solution to the equations.

4 Generic Results of Thermoacoustic Model With Fully Pre-Mixed Flame

Prior to application of the model to the actual gas turbine geometry, the model was run for some simple ductwork configurations in order to establish generic characteristics which would apply to a wide range of situations. In particular, attempts were made to establish how frequency and damping of the acoustic modes were affected by operational and geometric parameters. Now, non-dimensional analysis leads to relationships of the general form

$$\omega L/c_c = \mathrm{func}_1(T_h/T_c, \xi\mathrm{Mn}, \psi, \tau, \mathrm{geometry}) \tag{8}$$

$$\zeta = \mathrm{func}_2(T_h/T_c, \xi\mathrm{Mn}, \psi, \tau, \mathrm{geometry}) \tag{9}$$

An alternative expression for a given geometry, when the time delay scales inversely with swirl vane exit velocity and when ψ is constant or depends on $\omega\Delta t$, is simply $\zeta = \mathrm{func}_3(T_h/T_c, \mathrm{Mn})$, i.e., damping depends only on Mach number and temperature ratio.

The purpose of the generic work was to establish the form of the functional relationships func_{1-3} for simple cases when the operating parameters were broadly similar to those expected in a typical gas turbine. The results, some of which are presented below, give a valuable insight into the trends and order of magnitude of effects to be expected.

4.1 Localized or Distributed Pressure Drops. Localized or distributed pressure losses in the ductwork do not generally affect the frequency but will contribute to acoustic damping. For example, for the case of isothermal flow at (low) Mach number Mn down a duct with a distributed steady flow pressure drop coefficient given by the expression by $\xi = \Delta P/{}^1\!/_2\rho U^2$, the damping of the mth acoustic mode is given by the expression

$$\zeta = (1/m\pi)\cdot\xi\mathrm{Mn}. \tag{10}$$

Similar expressions can be derived for cases with concentrated pressure drops; however, in these cases the damping depends crucially on the position of the pressure losses compared to velocity mode shape, coincidence with velocity nodes leading to zero damping.

4.2 Cooling Air Holes. The effects of distributed cooling air flows, whether the flow is in or out, are quite similar to the effects of pressure drops in that they add to the system damping. Estimates of the resulting damping can be obtained from the expression

$$\zeta = (A_{\mathrm{cool}}/A)/(2\pi m \mathrm{Mn}_{\mathrm{cool}}), \tag{11}$$

where A_{cool} is the effective area of the cooling holes, A the main duct area, and $\mathrm{Mn}_{\mathrm{cool}}$ the Mach number of the air through the cooling holes calculated from the pressure drop.

4.3 Flame Position, Temperature Ratio, and Time Delay. The effect of a flame at various positions in a constant area duct, with and without velocity coupling (i.e., $\psi=1$ and $\psi=0$), and with and without time delay has been examined in some detail. This is a situation similar to Rayleigh's original experiment where conditions for thermoacoustic instability were demonstrated for the first time and the criteria for instability first established [15]. Some of the more important generic results are given below.

1 For the small time delays Δt, more accurately for small phase lags $\omega\Delta t$, frequency is independent of Δt but is a function of geometry, temperature ratio, and also flame coupling coefficient ψ. Therefore, the frequency cannot be calculated simply by reference to the passive acoustics of the hot and cold zones and account must be taken of the form of the flame/inlet velocity coupling as pointed out by Dowling [7]. Conversely, examination of the experimental trends in frequency can lead to an understanding of the extent of this coupling. Frequency is affected however when $\cos(\omega\Delta t)$ is not close to unity which will be true in most cases of interest.

2 With no time delay or no coupling or no temperature rise, there is neither damping nor amplification due to the flame. Furthermore, for small phase lags and moderate temperature rises $\Delta T/T_c$, a maximum value for damping (either positive or negative) can be shown to be of the order of $1/4\omega\Delta t\cdot\psi\cdot\Delta T/T_c$.

3 For the case of the fundamental $m=1$ mode, damping is negative only when the flame lies upstream of a critical point in the duct which depends on $\Delta T/T_c$. This point moves upstream from the mid point as $\Delta T/T_c$ increases from zero.

4 For higher order modes, as $\Delta T/T_c$ is increased the behavior of the damping ratio is quite complicated, in some cases showing switches from negative values (instability) to positive values (stability) over short ranges of temperature ratio. This behavior, resulting from changes in velocity and pressure mode shapes, can be very important and may be a reason why dynamic pressures are so load sensitive.

5 Effects of Fuel Supply on Stability

5.1 Effect of Pressure Fluctuations on Heat Release and Damping. For simplicity, the model used has so far assumed that the air and fuel are fully pre-mixed, implying that the air and fuel supply impedances are matched in some way. It is instructive to challenge this assumption and consider the effect of varying the fuel supply impedance independently. In particular, we will examine how the overall system damping is affected by variations in fuel mass flow and heat release in response to fluctuations in local pressure p_2'. These effects are characterized in terms of the variable $\varphi(\omega)$ defined by the equation $\varphi(\omega)\equiv Q'/p_2'c$.

5.2 Fuel Supply Impedance. Consider the simple case of a flame in the middle of a duct with no pressure losses and with boundary conditions $p_1'=p_4'=0$. It can be shown that for the case when temperature ratio is close to unity, for odd numbered modes $m=1,3\ldots$, damping ratio is directly proportional to $\varphi(\omega)$ and inversely proportional to m

$$\zeta=-(\gamma-1)\varphi(\omega)/\pi m. \tag{12}$$

Further analysis shows that $\varphi(\omega)$ is the product of three terms: (a) the ratio of fuel supply area at exit to the cold duct area; (b) a non-dimensional thermodynamic factor $(\Delta H/c_c c_f)$ that depends only on the calorific value of the fuel and the speeds of sound c_c and c_f in the fuel and air; and (c) the impedance Z_f of the fuel supply lines, so that

$$\varphi(\omega)=(A_f/A_c)(\Delta H/c_c c_f)\cdot(1/Z_f(\omega)). \tag{13}$$

Now in general although A_f/A_c is small, say of order 0.01, $\Delta H/c_c c_f$, is of order 200 for a machine operating on natural gas, so the value of $\varphi(\omega)$ will be high provided that $1/Z_f$ is large. Damping rather than amplification always occurs since $1/Z_f$ must have a negative real part related to the overall pressure drops ΔP_f in the fuel supply system. Furthermore, if the fuel supply natural frequency coincides with the combustion system natural frequency, then $|Z_f|$ will have a minimum value of the order of $\xi \mathrm{Mn}_f$, where ξ is the steady flow pressure loss coefficient for the fuel supply system, and Mn_f the Mach number in the fuel supply delivery line (assumed to be of constant area). In this case, a simple, approximate expression for the maximum effect of the fuel supply system on the overall damping can be derived as follows:

$$\zeta_{\max}\approx(\gamma/4m)\cdot\mathrm{Mn}_c\cdot(\Delta T/T_c)\cdot(P_c/\Delta P_f). \tag{14}$$

That is, small values of fuel supply pressure drop lead to very large (and positive) damping effects if the fuel supply and overall system natural frequencies coincide.

5.3 Example—V94.2 Fuel Supply. A representation of the acoustics of the V94.2 fuel supply system has been modeled as a series of pipes, volumes, and pressure drops assuming approximate design flow conditions. In the absence of a flame, the compliance $1/Z_f$ of the system describes a clockwise loop in the complex plane (Fig. 3) as frequency increases, $\mathrm{Re}(1/Z_f)$ always remaining negative, i.e., damping is always assured. The geometry has been adjusted so that resonance, indicated by the symbol at the position of minimum impedance, occurs at 77 Hz, coinciding with the known V94.2 humming frequency. The relatively high pressure drops in the fuel supply prevent very low impedances, but nevertheless the overall effect on the combustion system is to give an exceedingly high potential for system damping due to the fuel supply, of the order 0.30 (30 percent) for the $m=4$ mode. This compares with typical values measured on operating machines of a few per cent (see later).

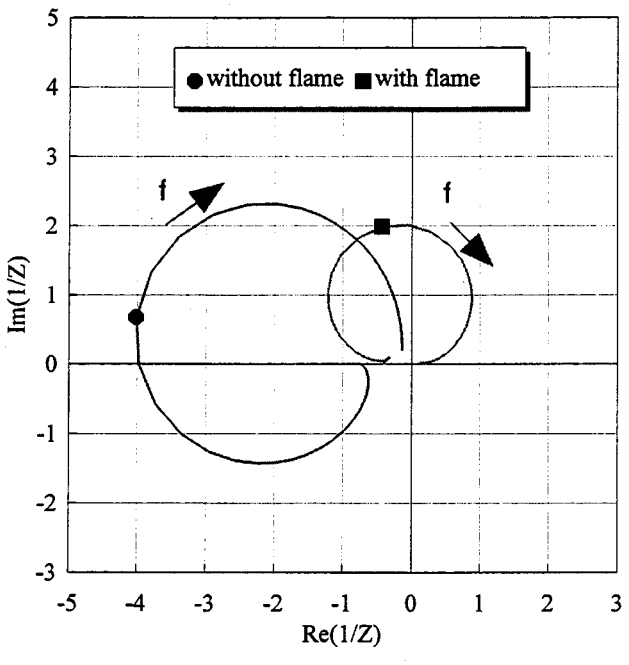


Fig. 3 Effective fuel supply compliance with and without flame

5.4 Modification of Fuel Supply Effects by the Flame. In practice, this beneficial effect of the fuel system will be modified by the flame itself. For example, if it is assumed that the same time delay function $\psi(\omega)$ applies as was the case for the pre-mixed supply, then the effective compliances shown in Fig. 3 will be multiplied by $\psi(\omega)$. This is illustrated in the second curve that represents points taken from the example above, but rotated by angles $(-j\omega\Delta t)$ and attenuated by a factor 0.5. It is clear that frequencies now exist when $\mathrm{Re}(1/Z_f)>0$, where coupling between the fuel supply and flame leads to instability. Inserting V94.2 values into Eqs. 13 and 14, the maximum value for the (now negative) damping ratio in this case is about 0.05 (5 percent), which is very significant in the light of the experimental values.

6 Determination of Flame Frequency Response Functions

6.1 Calculation of Time Delay Using Steady CFD. When applying the model to an actual machine, the major unknown is the behavior of the flame and in particular its frequency response function, characterized initially by a convective time delay Δt as suggested by previous authors (see Section 2). To obtain an estimate of Δt for use in the stability model, a CFD analysis was used. The detailed burner geometry was measured and representative inlet flow rates, pressures and temperatures for air and gas were calculated. An axisymmetric CFD model of a single V94.2 burner was then developed using a commercial code CFX4. The four individual pilot jets were modeled as a single annular jet. The combustion was modeled using an ''eddy break up'' model [16]. This assumes that the reaction rate is determined by how effectively the turbulence in the flow mixes fuel, oxidant, and hot products together, and that this mixing is performed by the smallest turbulent eddies. The effect of a finite chemical reaction rate is

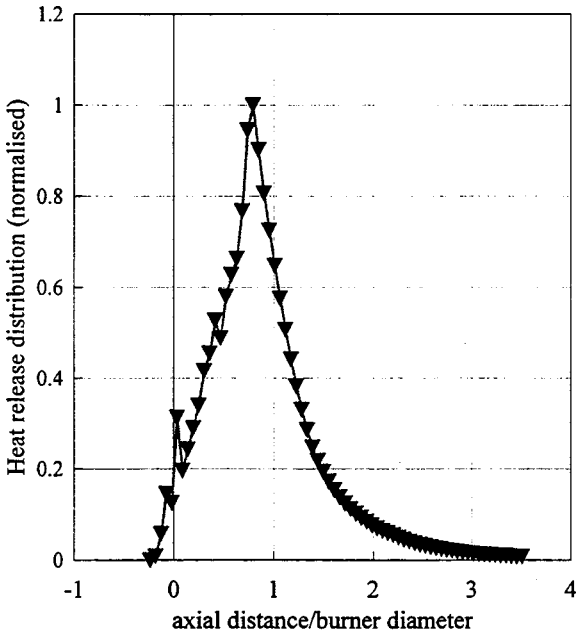


Fig. 4 Distribution of heat release in flame

included in the model following the approach taken by Bakke and Hjertager [17]. The ratio of turbulent eddy timescale to the chemical reaction timescale (the Damkohler number) is estimated for each grid cell and the combustion is suppressed locally if the ratio falls below a specified value (0.001). The Damkohler number has a strong dependence on temperature, so this cut-off effectively prevents combustion of the pre-mixed gas until its temperature is raised by mixing with the recirculated hot products. The values used to define the chemical reaction timescale were based on those given for a single step methane reaction by Westbrook and Dryer [18], but as later corrected and modified by Coffee [19].

The local heat generation rate was calculated within the CFD program by a user-supplied routine and this rate integrated radially and circumferentially to obtain the axial variation and hence identify the axial position of the maximum heat generation rate. A sample result for this variation is given in Fig. 4. The average time taken for fluid particles to convect from the fuel inlet to this position was then estimated by a separate post-process run of the particle tracking model in CFX. This time was taken as the time delay Δt to use in the stability model. This single estimate of Δt is a simplification for modeling purposes, since there is in fact a range around this value (say ± 20 percent), depending on the origin of the fluid particles. Note that the time delay of about 6 ms causes phase shifts between heat release and velocity of about 5 radians and, according to the theoretical model, this will cause significant shifts in frequency as well as affecting damping.

By changing input values, the steady state CFD model was used to examine the sensitivity of the time delay to changes in input parameters such as compressor delivery temperature, fuel mass fraction and flow rate (and hence stoichiometry) and these sensitivities were included in the stability model.

6.2 Flame Frequency Response Function From Transient CFD. It is recognized that the approach based on a simple time delay model calculated using steady CFD, though having the benefit of simplicity, is only an approximation even for the fully pre-mixed case and that in many cases the full frequency response function $\psi(\omega)$ (Eq. 5) should be evaluated. However, trials using transient CFD (not reported in detail here), where the volume flow of the pre-mixed gas was changed in a step-like fashion and the time variation in the overall heat release calculated, show that for this burner at least the approximation is a reasonable working assumption close to the frequencies of interest. In particular, at these frequencies the time delays calculated by the two methods are in close agreement though some attenuation of the quasi-steady response occurs, i.e., $|\psi|<1.0$.

As well as dealing with the fully pre-mixed, constant stoichiometry case, however, transient CFD can be used to calculate a second, quite different FFRF where the fuel supply is varied at constant air flow. (A third FFRF, where fuel supply is fixed and air supply is varied can also be evaluated but, for small disturbances, this is linearly related to the first two.) The FFRFs have different shapes but, over the limited frequency range of interest, can again be approximated by a time delay with Δt of between 6 and 7 ms and amplitude $|\psi|$ between 0.5 and 0.8.

7 Measurements on Operational Gas Turbines

One of the difficulties facing gas turbine operators when trying to understand the characteristics of combustion instability is taking detailed measurements inside the ducts of the machine. At best, operating machines will have a small number of dynamic pressure transducers somewhere inside the combustion chamber, at worst only external casing vibration transducers. Furthermore, all information must be obtained with the machine burners operating in a stable condition since operating beyond the stability threshold for extended periods not only results in damage to gas turbine internals but also violates the small perturbation analysis used in the theory. Fortunately, even in these stable conditions, peaks in the acoustic and vibration power spectra which might ultimately be associated with unstable behavior are easily identified; these peaks are caused by the excitation of the acoustic resonances inside the combustion chamber by broad-band combustion or flow noise. For these reasons, the basis for comparison between theory and measurement was chosen to be (a) the frequency ω_0 of peaks in the vibration or acoustic power spectra; (b) their half-power bandwidths $\Delta\omega$. From this information, the damping ratio $\zeta=\Delta\omega/2\omega_0$ was then derived.

7.1 Instrumentation and Data Recording. To determine these parameters, a transportable PC-based, 16 channel data acquisition system was installed on three machines GT01, GT02 and GT03 at two CCGT stations in the UK and left to capture data automatically for periods of up to six weeks. The machines were nominally identical but some minor modifications had been carried out in the past that might have had an influence on humming. The system could be controlled either locally or remotely from any convenient site equipped with a PC with suitable software and a modem link. Once set up, sampling and analysis would continue automatically until interrupted.

Dynamic pressures signals (where available), compressor casing vibration signals, load, IGV position and other operating parameters were input to the system from the stations' own vibration monitoring systems. The data was captured using the DASYLab (Data Acquisition System Laboratory) software suite of programs installed on the data acquisition system, the worksheet used being configured to capture the signals and calculate and store to disk FFT amplitude spectra. A very fine frequency resolution was chosen so that damping values down to about 0.001 (0.1 percent) could be resolved.

7.2 Data Processing to Obtain Frequency and Damping. The DASYLab system was programmed to carry out a fast Fourier transform of the pressure or vibration, signals $y(t)$ and stored the data as a series of values of pressure or vibration power spectral density at intervals equal to the analysis bandwidth. A MATLAB program was then used off-line to take the DASYLab data and calculate values ζ and ω_0 to best-fit a curve of the form

$$y(\omega)^2=B(\omega_0)/\{(1-(\omega/\omega_0)^2)^2+(2\zeta\omega/\omega_0)^2\} \qquad (15)$$

to the experimental data close to the resonant peak. Here B represents a locally constant amplitude (Pa or mm/s). The overall

energy y^2_{rms} associated with the resonance was also calculated by first converting the $y(\omega)$ values into a power spectral density and integrating over a suitable frequency range.

7.3 Preliminary Measurements of Dynamic Pressures. Some preliminary tests were carried out on GT01 over a period when the machine was operating mainly at base load (though some load swings took place), often in cold weather, with the objective of establishing typical levels of dynamic pressure in the hot zone immediately above the flame, and correlations with load and inlet temperature. Signals were band-passed filtered and the rms levels within the band 60 Hz to 90 Hz obtained. The results, presented as long term averages versus load, are shown in Fig. 5 at 1 MW intervals. It can be seen that there is a gradual rise in dynamic pressure from low load up to about 150 MW, followed by a rapid rise towards 10 mbar at the maximum load obtainable during this period, 169 MW. Attempts to find a second independent parameter (e.g., inlet temperature) against which to correlate were unsuccessful due to the scatter in the data. However, it was noted that the behavior of the data was similar to that expected from a resonant system excited by a roughly constant level of broad band noise but with a damping which declined in an approximately linear fashion towards zero at a load of about 173 MW (see fitted curve). This was an indication that variable damping was playing a role in the problem.

7.4 Measurements of Frequency and Damping During a Short Load Swing on GT03. In the absence of dynamic pressure, measurements of compressor casing vibration were taken on this machine during various phases of a load swing (see Fig. 6): on load; load drop; low load; load up; humming (during load up); load trim. Humming was thought to be a state where the gas turbine operated beyond the combustion stability limit when amplitudes were governed by non-linear effects. The machine became stable when the load was trimmed by slightly reducing the gas turbine exhaust temperature. The data points plotted are times at which the vibration data was analyzed.

Figure 7 shows how the frequency of vibration varied during the load swing. A trend upwards as load increased and humming commences can be observed. It is clear also that the on-load and load-trim data points are not part of the same family as the other data, with lower values of frequency (about 76.5 Hz) compared to 77.5 Hz at a load of 148 MW with humming. Since part of the reason for higher frequencies is higher hot gas temperatures, it is therefore possible that these temperatures were slightly but abnormally high on GT03 and this is likely to be a factor in inducing humming.

Figure 8 is perhaps the most significant result in that it demonstrates most clearly that increased vibration (and dynamic pressures) at high loads are caused primarily by a gradual decrease in damping. (Note that the lowest levels observed during humming, 0.001 or 0.1 percent of the critical damping, are the smallest that can be calculated using the analysis bandwidth of 0.125 Hz). Not

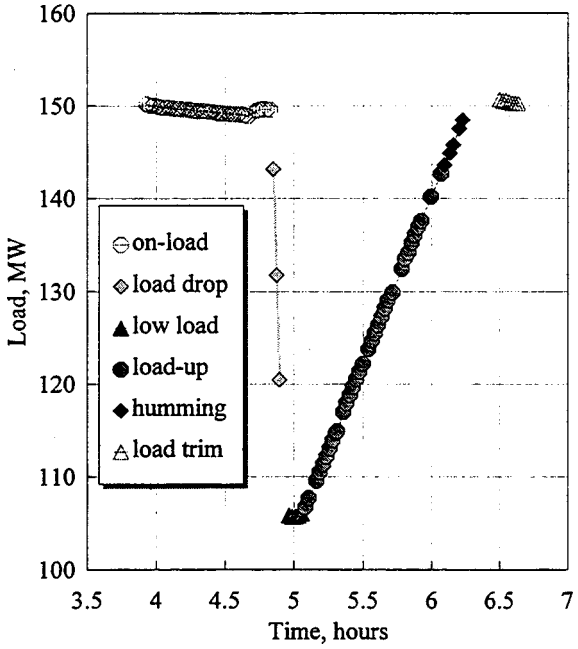


Fig. 6 Gas turbine GT03 load swing profile

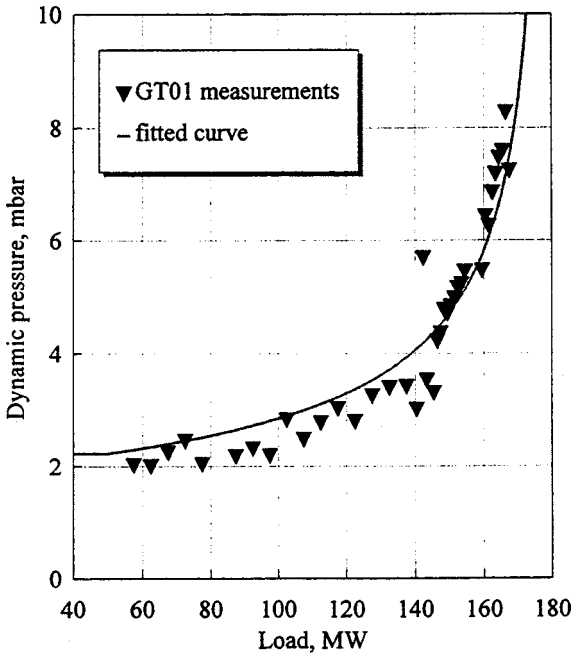


Fig. 5 Long term average dynamic pressures on GT01

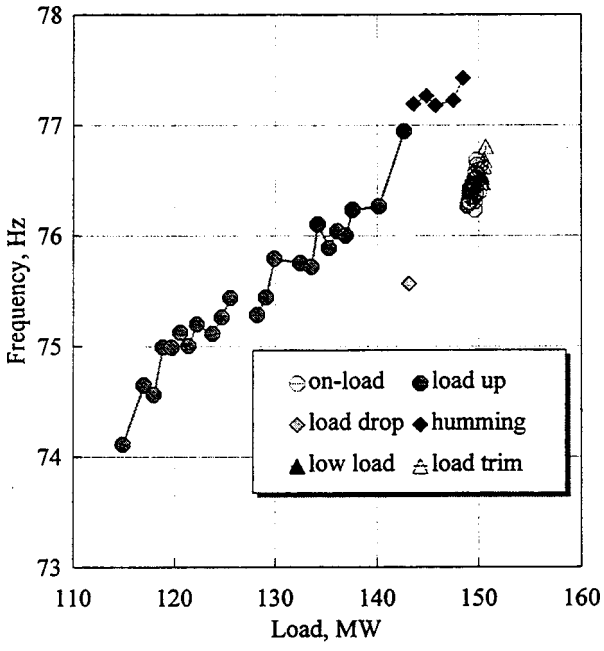


Fig. 7 Variation of frequency with load during GT03 load swing

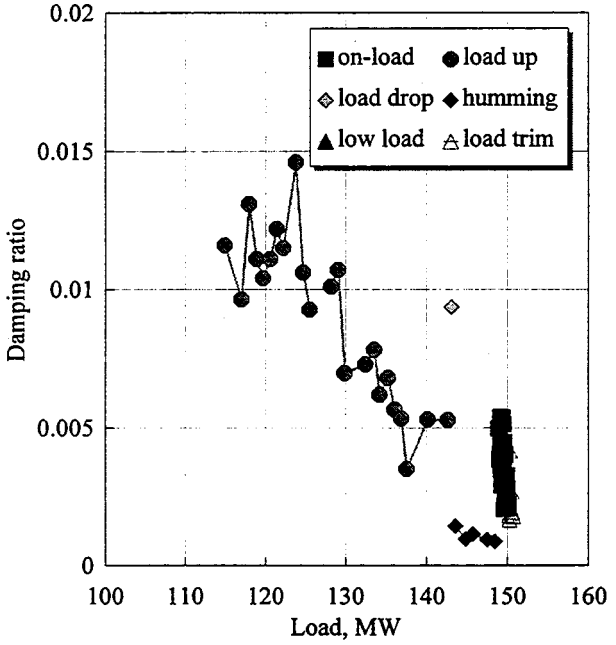


Fig. 8 Variation of damping with load during load swing on GT03

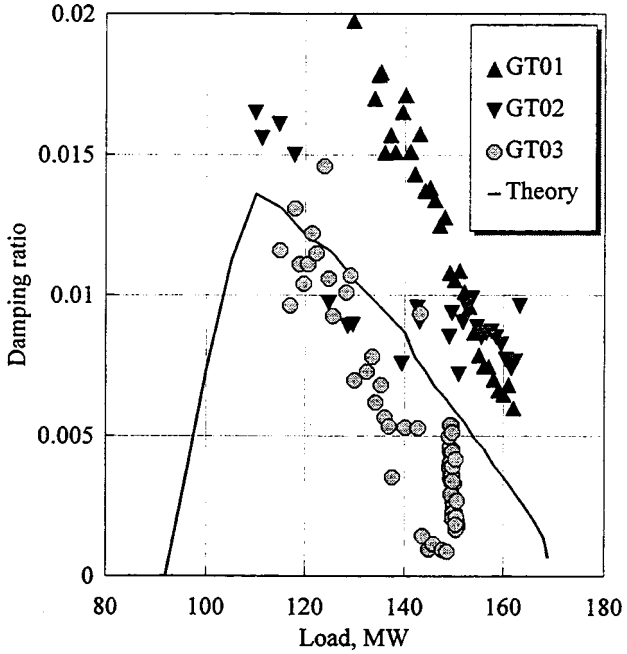


Fig. 10 All machines, variation of damping with load

shown are the calculated values of amplitude $B(\omega_0)$ (Eq. 15) which are measures of the level of broad band turbulence at the frequencies close to resonance; a level of B^2 proportional to load might be expected but there was little evidence of this.

7.5 Results From Extended Periods of Operation of GT01 and GT02. A similar analysis was carried out on gas turbines GT01 and GT02, making use of either fitted dynamic pressure transducers or the compressor bearing casing vibration signals. During these periods of operation, lasting four or six weeks, there were few significant load swings, though the load varied in sympathy with inlet temperature. The data was characterized by considerable scatter of up to ± 1 Hz in frequency and ± 0.005 in damping, due probably to changes in other independent parameters during the course of the tests. Averaged data has been calculated therefore and this is shown in Figs. 9 and 10.

On GT01, frequencies derived from dynamic pressure signals, were close to those of GT03. Like GT03, damping showed a clear trend downwards but the level was much higher and consistent with a stability limit about 20 MW higher than GT03. GT02's frequencies, derived from vibration measurements, were quite different however and trends in damping not so clear. It is thought that the reason for this is the distortion of the acoustic resonance by the presence of a local mechanical resonance with a similar, fixed frequency and difficulties with the curve fitting algorithms.

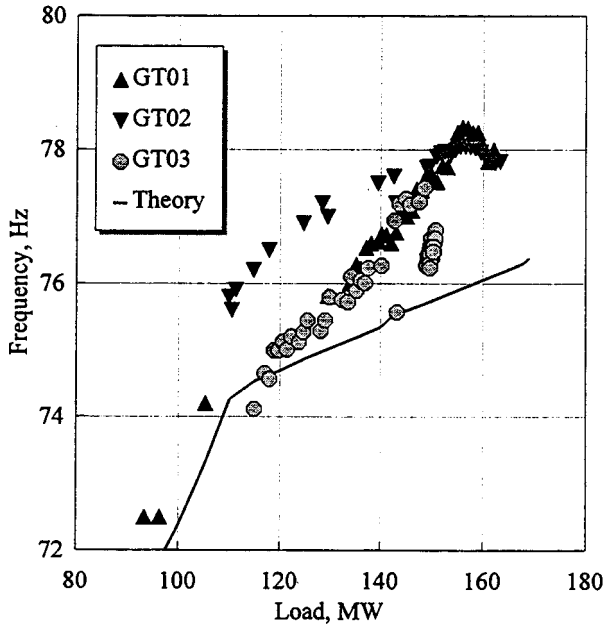


Fig. 9 All machines, variation of frequency with load

8 Comparison Between Measurements and the Thermo-Acoustic Model

During a load up of a V94.2, at loads above about 110 MW the compressor inlet guide vanes are continuously opened and fuel flow adjusted to maintain a constant turbine exhaust temperature. The non-dimensional parameters governing stability, i.e., temperature ratio, excess air ratio and compressor discharge Mach number, remain almost constant at values of about 2.4, 2.2, and 0.06, respectively. The increase in load arises from an increase in compressor mass flow and discharge pressure. To first order, it would be expected from Eq. 8 that frequency would rise in proportion to c_c but that damping would remain fixed. This is not the case in practice and second order effects take place, necessitating a more careful examination of how the operating parameters vary during a load up. This was achieved by a combination of overall gas turbine performance data measurement and thermodynamic calculations, primarily of flame temperature.

The thermo-acoustic model was run with these parameters as input and with the time delay estimates from the CFD modeling and the results compared with the experimental measurements of frequency and damping. In Figs. 9 and 10 the experimental results for all three turbines are compared with the model predictions for the 4th acoustic mode, this mode having frequencies closest to the experimental ones and with trends in damping closest to those observed. (Note that the kink in the theoretical curve correspond to a change in the way in which the machine is controlled.) Allowing for the variability in the experimental results, it can be

seen that the model predictions agree reasonably well with measurements, in particular both the increase in frequency and decrease in damping with load are reproduced. However, to obtain this degree of agreement some tuning of the model was needed. The effective acoustic lengths of the upstream and downstream ducts which were initially estimated as 7.7 m and 8.0 m had to be reduced by 5 percent and 6 percent respectively and the time delay given by the CFD analysis was increased by 15 percent. This level of adjustment was not unexpected, since the acoustic lengths cannot be estimated very accurately, (the real system not being just two simple ducts), and the time delay from the CFD modeling is only an approximate value. Since the predicted results are quite sensitive to the ratio of acoustic lengths and these will always need some refinement, the simple acoustic model used here should be seen more as a tool for understanding the plant behavior rather than as a truly predictive tool, at least for this model of gas turbine.

A sensitivity analysis showed that even small changes in the calculated values of volumetric flow rate, Mach number, temperature ratio, and time delay had considerable influences on the predictions of both the frequency and damping. Therefore, the results are also sensitive to the thermodynamic model and in particular the method of calculating the effective flame temperature from the gas turbine exhaust temperatures. It is thought that these effects and measurement uncertainties in the exhaust temperatures were responsible for the premature humming of GT03.

9 Conclusions

A theoretical model has been developed which has been used to derive some generic characteristics of combustion instability in terms of damping ratio. It is based on an acoustic model for the combustion chamber ducts and a frequency response function model for the flame.

A CFD model of flow and combustion in a Siemens V94.2 burner has been used to calculate a convective time delay which can be used to derive an approximation to the frequency response function.

Some generic aspects of the effects of fuel supply impedance have been explored. It has been shown that a fuel supply resonating at a frequency close to the overall system natural frequency can have a considerable influence on stability but time delays in the flame must be included when calculating these effects.

Quantitative measures of the stability of combustion on three operating gas turbines have been obtained by measurements of the frequency and bandwidth of peaks in the vibration or dynamic pressure spectra. Values for damping ratio have been derived from these measurements. As load increased, frequency increased and damping decreased steadily, the latter being the main reason for the rise in the signal amplitudes. Damping reached zero on one machine and a true instability occurred.

The theory has been applied to the operating gas turbines using measured or calculated values for the performance characteristics and the results of the CFD calculations as input. The results show reasonable agreement but also a very large sensitivity to the calculated values of volumetric flow rate and flame temperature as well as some of the geometric parameters. These sensitivities, combined with uncertainties in the calculation of some of the performance parameters on operating machines, may be one of the reasons why nominally identical machines have different stability limits. They also limit the fully predictive capabilities of the model.

Nomenclature

A_c, A_h = cold and hot duct areas
A_{cool} = area of cooling holes
$B(\omega_0)$ = amplitudes of lines in FFT
c_v, c_p = specific heats of air
c = velocity of sound
ΔH = fuel calorific value
j = $\sqrt{-1}$
L_c, L_h = cold and hot duct lengths
m = mode number, $m=1,2\ldots$
Mn = Mach number
$\mathbf{M}(r)$ = coefficient matrix
p, P = pressure; mean pressure
ΔP = mean pressure drop
Q = heat release rate per unit cross sectional area
r = variation with time is $\exp(rt)$, $r \equiv s + j\omega$
s = $\mathrm{Re}(r)$
t = time
Δt = time delay
T = temperature
ΔT = $T_h - T_c$
u, U = velocity; mean velocity
x = flow direction
y = vibration or pressure signal, Pa or mm/s
γ = ratio of specific heats
Z = acoustic impedance, $p'/\rho c u'$
ζ = damping ratio, $-\mathrm{Re}(r)/\mathrm{Im}(r)$
ξ = pressure drop coefficient, $\Delta P / {}^1/_2 \rho U^2$
ρ = gas density
τ = non-dimensional time delay, $c_c \Delta t / L_c$
$\varphi(\omega)$ = frequency response function between heat release and acoustic pressure, $\varphi(\omega) \equiv Q'/c_c p_2'$
$\psi(\omega)$ = flame frequency response function (FFRF) relating heat release and acoustic velocity, $\psi(\omega) \cdot (\Delta T/T_c)/(\gamma - 1) \equiv Q'/\rho_c c_c u_2'$
ω = angular frequency, $\mathrm{Im}(r)$

Subscripts

1,2,3,4 = compressor discharge, burner inlet, burner outlet, combustion chamber exit locations
a, f = air, fuel
c, h, cool = cold duct, hot duct, cooling holes
s = swirl vane exit conditions

References

[1] Richards, G. A., Morris, G. J., Shaw, D. W., Keeley, S. A., and Welter, M. J., 1993, ''Thermal Pulse Combustion,'' Combust. Sci. Technol., **94**, pp. 57–85.
[2] Janus, M. C., Richards, G. A., Yip, M. J., and Robey, E. H., 1997, ''Effects of Ambient Conditions and Fuel Composition on Combustion Stability,'' ASME Paper 97-GT-266.
[3] Smith, C. E., and Leonard, A. D., 1997, ''CFD Modeling of Combustion Instability in Pre-Mixed Axisymmetric Combustors,'' ASME Paper 97-GT-305.
[4] Bohn, D., Deutsch, G., and Kruger, U., 1996, ''Numerical Prediction of the Dynamic Behavior of Pre-Mixed Flames,'' ASME Paper 96-GT-133.
[5] Bohn, D., Yuhong, L., Matouschek, G., and Kruger, U., 1997, ''Numerical Prediction of the Dynamic Behavior of Pre-Mixed Flames Using Systematically Reduced Multi-Step Reaction Mechanisms,'' ASME Paper 97-GT-265.
[6] Liu, R., and McGuirk, J. J., 1995, ''Prediction of Combustion Induced Oscillations Using a Pressure–Correction Method,'' ASME Paper 95-GT-336.
[7] Dowling, A. P., 1995, ''The Calculation of Thermoacoustic Oscillations,'' J. Sound Vib., **180**, Pt. 4, pp. 557–581.
[8] Chadha, J. A., Hobson, D. E., Marshall, A., and Wilkinson, D. H., 1980, ''Acoustic Source Properties of Governor Valves,'' in *Flow Induced Vibrations of Power Plant Components*, M. K. Au-Yang, ed., ASME PVP-41.
[9] Buchner, H., Hirsch, C., and Leuckel, W., 1993, ''Experimental Investigations on the Dynamics of Pulsated Pre-Mixed Axial Jet Flames,'' Combust. Sci. Technol., **94**, pp. 219–228.
[10] Keller, J. O., Bramlette, T. T., Dec, J. E., and Westbrook, C. K., 1989, ''Pulse Combustion: The Importance of Characteristic Times,'' Combust. Flame, **75**, pp. 33–44.
[11] Bloxsidge, G. J., Dowling, A. P., and Langhorne, P. J., 1988, ''Reheat Buzz: An Acoustically Coupled Combustion Instability. Part 2—Theory,'' J. Fluid Mech., **193**, pp. 445–473.
[12] Lieuwen, T., and Zinn, B. T., 1998, ''Theoretical Investigation of Combustion Instability Mechanisms in Lean Pre-Mixed Gas Turbines,'' AIAA Paper AIAA-98-0641.
[13] Bohn, D., and Deuker, E., 1993, ''An Acoustic Model to Predict Combustion Driven Oscillations,'' *20th Int. Congress on Combustion Engines*, London.
[14] Becker, B., Berenbrink, B. P., and Brandner, H., 1986, ''Premixing Gas and

Air to Reduce NOx Emissions With Existing Proven Gas Turbine Combustion Chambers,'' ASME Paper 86-GT-157.
[15] Rayleigh, J. W. S., 1945, *The Theory of Sound*, Vol. II, pp. 224–227, Dover, New York,
[16] Mason, H. B., and Spalding, D. B., 1973, ''Prediction of Reaction-Rates in Turbulent Pre-Mixed Boundary Layer Flows,'' *Combustion Institute European Symposium*, E. J. Weinberg, ed.
[17] Bakke, J. R., and Hjertager, B. H., 1987, ''The Effect of Explosion Venting in Empty Vessels,'' Int. J. Numer. Methods Eng., **24**, pp. 129–140.
[18] Westbrook, C. K., and Dryer, E. L., 1981, ''Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames,'' Combust. Sci. Technol., **27**, pp. 31–43.
[19] Coffee, T. P., 1985, ''On Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames,'' Combust. Sci. Technol., **43**, pp. 333–339.

Natural Gas Fired Combined Cycles With Low CO_2 Emissions

Paolo Chiesa
e-mail: pchiesa@clausius.energ.polimi.it

Stefano Consonni
e-mail: stefcons@axp7000.cdc.polimi.it

Dipartimento di Energetica
Politecnico di Milano
P.zza Leonardo da Vinci, 32
20133 Milan, Italy

This paper assesses performances and economic viability of CO_2 removal by chemical absorption from the flue gases of natural gas-fired Combined Cycles, more specifically for two configurations: one where CO_2 is removed ahead of the stack without modifying the power cycle; the other where part of the flue gases is recirculated to the gas turbine, thereby reducing the flow to be treated by chemical absorption. In both cases sequestered CO_2 is made available at conditions suitable to storage into deep oceanic waters. Performances and cost of electricity are evaluated for systems based on large, heavy-duty turbines representative of state-of-the-art ''FA'' technology. Carbon sequestration reduces net plant efficiency and power output by about 10 percent and increases the cost of electricity from 36 to about 50 mills/kWh. Flue gas recirculation warrants slightly higher efficiencies and lower costs. CO_2 removal is eventually compared with other strategies for the reduction of CO_2 emissions, like switching existing coal-fired steam plants to natural gas or replacing existing steam plants with conventional CCs. At current fuel prices the latter appears the option of choice, with a cost of about $25 per tonn of avoided CO_2 emission. [S0742-4795(00)02803-9]

Background and Scope

Fossil fuels are likely to remain the main source of primary energy for a long time. The CO_2 emissions due to their use can be reduced either by higher efficiencies or by carbon sequestration. The latter can be accomplished according to three basic concepts:

1 remove CO_2 from combustion gases, with no or minimal modifications to existing plants
2 restructure the power cycle so that carbon (as CO_2) can be removed along the cycle itself
3 ''decarbonize'' the fuel, i.e., remove carbon (also in this case as CO_2) to produce hydrogen

This paper concentrates on natural gas-fired Combined Cycles (CC) where the implementation of concept (1) lowers specific CO_2 emissions (kg_{CO_2}/kWh) to about 10 percent of those of conventional CCs.

One could argue that, being the fossil fuel that generates the least CO_2, natural gas should be given the lowest priority for carbon sequestration. Indeed this is not the case, because the most relevant parameters for the allocation of plants with low CO_2 emissions are the cost per output unit (in our case, the cost per net kWh generated) and the cost per unit of avoided CO_2 emission. Should those costs be lower for natural gas-fired systems, the optimal arrangement of the power generating system under a constraint on total CO_2 emissions may well consist of a combination of natural gas-fired plants with nearly zero CO_2 emissions and plants fired with more carbon intensive fuels (like coal) with relatively large CO_2 emissions. The final section of the paper gives quantitative indications for the selection of the optimal strategy to reduce the CO_2 emissions of an electric power generation system.

Systems Considered in the Paper

Figure 1 provides a conceptual overview of the two configurations considered in the paper. Both of them are built around the same 230 MW class heavy-duty gas turbine, with a three-pressure-level, reheat steam cycle meeting the standards of state-of-the-art CC technology. Sequestered CO_2 is made available as liquid at 80 bar, thereby warranting either reduced transport costs via pipeline or the discharge into oceanic water at a depth between 1000 and 1500 m, where the probability of CO_2 ascending to the surface appears remote [1].

Scheme A is just a conventional CC with the addition of CO_2 chemical absorption ahead of the stack. Given the incompatibility between monoethanolamine (MEA) and oxygen-rich gases [2], we have considered an aqueous solution of diethanolamine (DEA), with 40 percent DEA by weight. Scheme B adds to Scheme A the recirculation of part of the exhaust gases to the gas turbine compressor (dashed line to compressor inlet). This recycle reduces the flow rate of exhaust gases to be treated by increasing their CO_2 concentration, whereby lower steam consumption in the stripper and smaller and cheaper equipment. Compared to conventional gas turbines, the working fluid is rich in CO_2 and low in O_2; this produces only minor consequences to the gas turbine operating conditions, because the relevant turbomachinery parameters (enthalpy drop, speed of sound, ratio of specific heats, ratio of volume flows) change by small amounts [3].

The ''base CC'' used as a reference is a conventional plant based on the same gas turbine assumed for the systems with CO_2 removal and with the same steam cycle parameters. The chemical absorption system conforms to standard ethanolamine plants [2].

The two schemes in Fig. 1 do not exhaust the possibilities for CO_2 sequestration. Alternative arrangements are as follows:

1 Semiclosed cycle with combustion into pure oxygen, whereby combustion products consist almost exclusively of CO_2 and H_2O, and CO_2 is separated by cooling the flue gases below the dew point. This scheme is an example of strategy (2) mentioned in the introduction (restructure the cycle). Despite the absence of a CO_2 removal section, the implementation of this concept is made difficult (and costly) by the need for an air separation unit and a complete redesign of the gas turbine. Analyses carried out by Chiesa and Lozza [4] and Chiesa, Consonni and Lozza [5] for coal fired plants show that when the gas cycle parameters are similar to those of conventional gas turbines (atmospheric pressure at compressor inlet, pressure ratio in the range 20–40, TIT around 1300°C), this scheme does not look particularly attractive; the same applies to natural gas fired systems [6]. Much better prospects are envisaged for the supercritical, regenerative, reheat

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-370. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

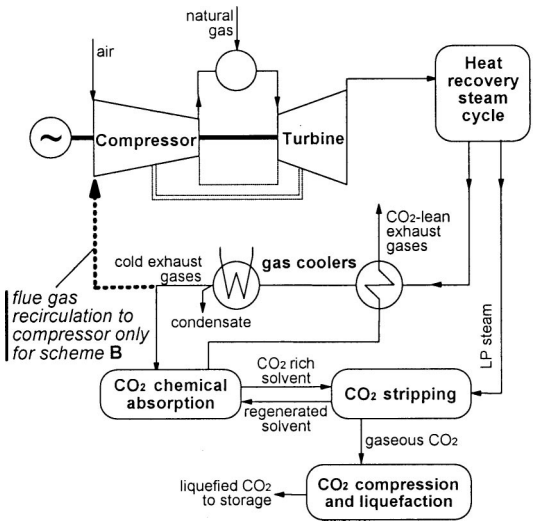


Fig. 1 Combined cycle configurations considered in the paper. The difference between schemes A and B lies in the recirculation of part of the flue gases to the gas turbine compressor.

cycle proposed by Mathieu and Nihart [7], where the working fluid (mostly CO_2) is compressed in the liquid state beyond 300 bar.

2 Fuel decarbonization by steam reforming or partial oxidation followed by shift and CO_2 chemical (or physical) absorption [8]. In this case natural gas is converted into an hydrogen-rich fuel carrying most of the original heating value. This fuel can be burnt either in a conventional CC or in the semiclosed H_2O cycle considered by Bannister et al. [9] and Kizuka et al. [10]. Being essentially CO_2-free, combustion products do not pose greenhouse gas issues.

These alternative configurations share the need for developing components new per se (e.g., the CO_2 turbine) or in their implementation to power plants (e.g., steam reformer relying on convective heat transfer). On the contrary, the two systems considered in this paper appear closer to commercial readiness because, even if chemical absorption is new to power plants, its introduction does not require any specific re-design.

Modelization and Operating Conditions

Efficiency and power output have been estimated by a computer code designed to evaluate heat and mass balances of complex gas-turbine-based power plants. The model, which is extensively described elsewhere [11–14], allows taking properly into account the implications posed by the DEA absorption system (export of steam to the stripper) and by the flue gas recirculation of scheme B (changes in the composition of the gas turbine working fluid). The chemical absorption system has been evaluated by the ASPEN code [15]. The use of two different codes has required some iteration to match the specifications of the absorption plant (stripper steam consumption) with those of the power plant (amount of CO_2 to be removed).

Table 1 summarizes the most relevant assumptions. Gas turbine air flow, turbine inlet temperature and pressure ratio are representative of state-of-the-art, large heavy-duties, i.e., what is commonly called ''FA'' technology. The three-pressure level reheat steam cycle conforms to the standard adopted for large CCs based on such ''FA'' turbines [16].

In Scheme B, the mass flow of flue gases recirculated to the compressor and the mass flow of fresh air are determined by imposing that the volume flow at the compressor inlet is the same of

Table 1 Main assumptions maintained throughout all calculations

GAS TURBINE CYCLE	
Pressure ratio	15.1
Turbine inlet temperature, °C	1316
O_2 content at combustor exit in schemes B, %	3
Volume flow at compressor inlet, m^3/s	517.3
The volume flow at compressor inlet corresponds to an inlet mass flow of 625 kg/s for an open-loop gas turbine	
STEAM CYCLE (3 pressure level + RH)	
Maximum steam evaporation pressure, bar	140
Steam temperature at turbine admission, °C	565
Condensation pressure, bar	0.05
CO_2 ABSORPTION AND COMPRESSION	
DEA concentration, % by weight	40
Max. CO_2 concentration at absorber inlet, mol_{CO2}/mol_{DEA}	0.15
Stripping pressure, bar	2
Minimum ΔT at regenerator exit, °C	10
Absorber and stripper packing height, m	4
Size of packing type (metallic HYPAK), mm	90
Maximum gas Δp in the absorber packing, kPa	4
Maximum gas Δp in the stripper packing, kPa	10
Number of intercoolers for CO_2 compression	3
Natural gas, molar composition %: CH_4=91.18, C_2H_6=4.41, C_3H_8=0.10, N_2=4.31; LHV: 46.30 MJ/kg; HHV: 51.31 MJ/kg	
CO_2 removal efficiency is adjusted to obtain a 90% reduction of specific emissions (kg_{CO2}/kWh).	

Scheme A (517.3 m^3/s, which is like saying that the gas turbine is the same) and the oxygen content at the combustor exit is 3 percent by volume. The assumptions for chemical absorption comply with those of previous studies [17–19] and with the practice reported by Kohl and Riesenfeld [2].

Figure 2 illustrates the configuration of the DEA absorption system. The CO_2-rich DEA solution exiting the absorber is preheated in a regenerative heat exchanger which cools the CO_2-lean solution, then it enters the stripper where medium-pressure steam provides the heat to reverse the CO_2-DEA reaction. Under the assumptions in Table 1, the design and operating conditions which match the requirements of the power cycle are those listed in

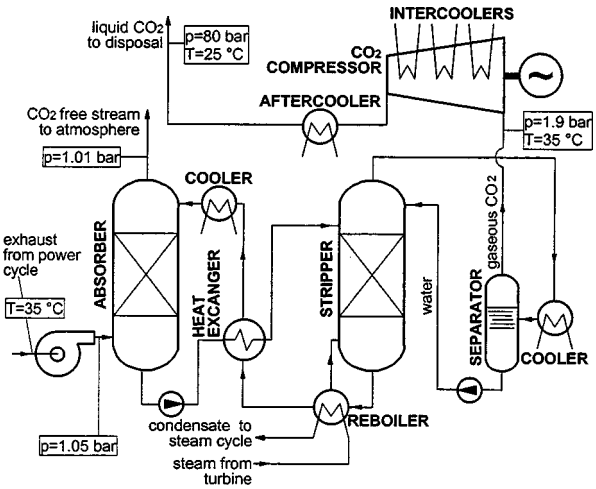


Fig. 2 Chemical absorption system considered in the paper. The configuration is representative of typical ethanolamine plants adopted in the chemical industry [2]

Table 2 Operating conditions of the DEA absorption system

	SCHEME A		SCHEME B	
ABSORBER	Inlet	Outlet	Inlet	Outlet
Gas flow rate at absorber inlet				
by mass, kg/s	627.7	602.9	283.8	256.9
by volume, m^3/s	538.7	558.9	240.9	244.0
Gas composition (mol fractions, %)				
Ar	0.91	0.92	0.95	0.99
CO_2	3.86	0.38	8.41	0.83
H_2O	5.55	7.87	5.60	9.30
N_2	76.72	77.69	80.19	83.81
O_2	12.97	13.13	4.85	5.06
CO_2 concentration, mol_{CO2}/mol_{DEA}	0.148	0.348	0.144	0.388
Gas temperature, °C	39.1	43.8	40.1	47.1
Solvent temperature, °C	35.0	42.6	35.0	42.6
Overall cross section, m^2	145.6		73.0	
STRIPPER				
Temperature of solvent at inlet, °C	101.1		99.5	
Reboiler heat duty, MW	110.5		97.5	
Steam bleeding pressure, bar	2.1		2.1	
Reflux ratio at stripper outlet, mol_{CO2}/mol_{steam}	1.15		1.28	
Overall cross section, m^2	16.3		14.0	
Solvent (40% DEA) mass flow rate, kg/s	1000		800	
CO_2 to disposal, kg/s	33.53		32.69	
Lean-rich heat exchanger duty, MW	215.5		163.7	
CO_2 absorption coolers' duty, MW	85.63		81.41	
CO_2 compressor power consumption, MW	9.66		9.42	
Heat released from inter and aftercooler, MW	18.99		18.40	
Pumps power consumption, MW	0.20		0.16	
Fan power consumption, MW	2.64		1.48	

Table 2.[1] Given the reference to actual, commercial equipment, operating conditions have been chosen in accordance to common industrial practice, without attempting the optimization of the integrated CC/DEA system[2]; performances and costs presented here might therefore be ameliorated.

Investment costs have been estimated by defining the cost scaling parameter, the base size and the base specific cost of each of the major subsystems of the plant, and then by estimating the actual cost according to a scale factor of 2/3 for subsystems subject to economies of scale, 1 for modular subsystems. Scaling parameters and base costs are reported in Table 3; the cost of the gas turbine needs no scaling because all plants are based on the same engine. The gas turbine and steam cycle investment costs interpret indications and trends reported in the literature for CCs centered around a large ''FA'' turbine; costs for absorption systems are in between those quoted by Smelser et al. [17], Hendricks [18], McMullen et al. [20] and Corti et al. [19].

Table 4 illustrates the assumptions adopted to estimate the cost of electricity (COE) [21,18,22]. The table includes also coal-fired plants because the search for an optimum strategy to reduce CO_2 emissions from power stations (see further) must necessarily consider the competition between coal and natural gas. ''IGCC'' and ''IGCC+shift'' refer to the Texaco coal gasification plants considered in Chiesa, Consonni and Lozza [5], the latter including shift reactors and physical absorption for the removal of CO_2 (see also Chiesa and Consonni [23]). The fuel costs of 3.3 and 1.7 \$/GJ[3] for natural gas and coal, respectively, are meant to represent medium term trends for the European market. The cost of CO_2 storage of 10 \$/tonn of CO_2 is in between the estimates for storage in deep oceanic waters via submerged pipeline and power plant placed by the sea shore (Fujioka et al. [24]).

Performances, Cost of Electricity, and Carbon Tax

Table 5 compares the most relevant performance indicators of schemes A and B with those of the base natural gas fired CC (NGFCC). Table 6 reports the investment costs obtained after scaling the base costs in Table 3 to the specific conditions in Table 5.

CO_2 removal reduces net power output by about 10 percent and efficiency by about 5 percentage points. In Scheme B, flue gas recirculation reduces gas turbine power output due to higher temperatures at compressor inlet, while lower heat for CO_2 removal gives larger steam cycle output; on the whole, power output is slightly less than with Scheme A, but efficiency is higher and investment costs are lower.

Table 7 shows that the cost of electricity rises from 36 mills/kWh to about 50 mills/kWh. Even if differences in investment cost and cost of electricity are small and yet considering the uncertainties of our analysis, the picture given by Tables 5 to 7 suggests that Scheme B with flue gas recirculation is more appealing than Scheme A, because it combines higher efficiencies, lower investment cost and lower cost of electricity.

[1]The design and operating conditions in Table 2 have been found by imposing that the amount of CO_2 removed lowers specific emissions (kg_{CO_2}/kWh) by 90 percent and that the concentration of CO_2 in the lean solvent entering the absorber is lower than 0.15 mol_{CO_2}/mol_{DEA}. The latter corresponds to usual DEA systems practice [2].

[2]Parameters subject to optimization are stripper pressure, CO_2 concentration in lean solvent at absorber inlet packing type and height of absorber and stripper, temperature differences in the regenerative heat exchanger, type and concentration of amine solution, etc. The optimization of some of these parameters is investigated in Corti et al. [19].

[3]Fuel specific costs, as well as efficiencies, are always referred to the fuel Lower Heating Value (LHV).

Table 3 Parameters adopted to scale investment costs. The cost C of an item with scale S different from the base size S_0 is evaluated by the relationship: $C=C_0(S/S_0)^f$. $f=1$ applies to modular subsystems for which there are no economies of scale. The data in the table are the same used to assess coal-fired systems in Chiesa et al. [5].

Component	Scale parameter	Base specific cost, c_0	Base size S_0	Scale factor, f
Gas turbine	Electric power	230 \$/kW	238 MW	-
Steam cycle	Electric power	585 \$/kW	200 MW	2/3
CO_2 compression system	Power consumed	400 \$/kW	30 MW	2/3
Chemical absorbers	Gas volumetric flow rate	52000 \$/($m^3 \cdot s^{-1}$)	60 m^3/s	1
Components of chemical absorption other than absorbers	Solvent mass flow rate	23000 \$/($kg \cdot s^{-1}$)	160 kg/s	1

Figure 3 reports the sensitivity of the marginal cost of electricity[4] of the technologies in Table 7 to a carbon tax proportional to the actual emissions of CO_2 to the atmosphere; given that the marginal costs of existing plants do not include amortization, the line for NGFSCs also gives the cost of switching existing coal fired plants to natural gas. With no carbon tax, cheap coal (1.7\$/GJ) makes existing pulverized coal steam cycles (PCSC) the cheapest option[5]; at a carbon tax of about 25 $\$/tonn_{CO_2}$ PCSCs leave the way to conventional (new) NGFCCs, while for a carbon tax close to 40 $\$/tonn_{CO_2}$ the cheapest option becomes the CC with CO_2 removal configured like Scheme B. In the absence of natural gas as an alternative to coal, PCSCs remain the cheapest option up to a carbon tax of about 70 \$/GJ, above which IGCCs with CO_2 removal become the option of choice at a cost of about 85 mills/kWh. It's worth noting that under no circumstance one should consider switching existing plants from coal to NG.

Figure 4 shows the competition between conventional NGFCCs and Scheme B as a function of natural gas price and carbon tax. In the leftward portion of the figure, conventional CCs give lower costs of electricity, the opposite in the remaining portion of the diagram. In the rightward region where Scheme B prevails, COE is almost independent of carbon tax because CO_2 emissions are almost zero. A carbon tax above 50 $\$/tonn_{CO_2}$ makes Scheme B more economic than conventional CCs for natural gas prices up to 10 \$/GJ.

The relative COE increase caused by the carbon tax that makes Scheme B competitive decreases with the price of natural gas: at present NG prices of 3.3 \$/GJ, CO_2 removal becomes attractive for a carbon tax which increases COE by 40 percent (from 36 to about 50 mills/kWh); at a price of 7 \$/GJ, CO_2 removal becomes attractive for a carbon tax which increases COE by about 25 percent (from 61 to about 77 mills/kWh). In other words, the higher the gas price the lower the relative impact on COE of CO_2 removal.

[4]''Marginal'' denotes the reference to existing power plants for which the additional costs incurred in the generation of one kWh are the ones in the two rightmost columns of the bottom line of Table 7. Being marginal, such costs do not include amortization. For new plants yet to be built marginal costs are total costs, i.e., they include amortization.

[5]The attribute ''existing'' is crucial to this statement. If one would consider new PCSCs, whereby marginal COE would include amortization, the ranking of PCSCs would change substantially.

Table 4 Assumptions for the estimation of the cost of electricity. Amortization is carried out in current dollars, whereby the carrying charge fraction for investment, property taxes and insurance varies from year to year (0.149 for the first year of operation year). O&M costs for conventional plants conform to the indications of EPRI [21]; for plants with CO_2 removal the same overall cost is charged to the reduced power output estimated for each scheme. The cost of consumables for CO_2 removal conforms to those quoted by Hendricks [18] and Doctor et al. [22].

	New plants					Existing steam cycle plants	
Plant configuration	CC	Scheme A	Scheme B	IGCC	IGCC + shift		
Fuel	NG	NG	NG	Coal	Coal	NG	Coal
Discount rate (current \$), %	10					-	
Inflation rate, %	2					-	
Income tax rate, %	40					-	
Plant life, years	25					-	
Tax recovery	straight-line depreciation over a 20 year period					-	
Property taxes and insurance	2% of plant cost per year, escalated for inflation					-	
Construction period, years	2	3	3	5	5	-	
Cost of fuel*, \$/GJ	3.3	3.3	3.3	1.7	1.7	3.3	1.7
Plant operation, h/y	7000					7000	
O&M*:							
fixed, % of plant cost per year	1.5	1.5	1.5	2.0	2.0	1.5	2.0
labor, \$/kW-year	4.00	4.44	4.50	5.75	6.35	5.75	5.75
consumables, mill\$/kWh power plant	0.40	0.44	0.45	1.00	1.18	0.40	2.50
CO_2 removal	-	0.75	0.75		0.10	-	-
Byproducts, mill\$/kWh	-	-	-	-0.50	-0.52		
Cost of CO_2 storage*, \$/tonn	-	10	10	-	10	-	-

Table 5 Overall performances of the plants considered in the paper. The "heat to cooling fluids" is the heat discharged at the condenser and the flue gas cooler.

Plant configuration	NGFCC	Scheme A	Scheme B
Fresh air at compressor inlet, kg/s	625.0	625.0	290.6
Recirculated flue gas, kg/s	0.0	0.0	314.2
Total mass flow at GT inlet, kg/s	625.0	625.0	604.8[1]
Temperature at turbine outlet, °C	610.6	610.6	622.2
Gas turbine power, MW	237.6	237.6	226.3
Steam turbine power, MW	139.8	116.3	121.6
CO_2 compression, MW	-	9.7	9.4
CO_2 removal and CC auxiliaries, MW	3.4	7.6	6.9
Net plant output, MW	373.2	336.6	331.6
Cycle thermal input (LHV), MW	665.3	665.3	649.1
Net plant efficiency, %	56.10	50.61	51.09
Heat to CO_2 separation process, MW	-	110.5	97.5
Heat to cooling fluids, MW	228.9	288.1	329.2
Removed CO_2, kg/s	0.00	33.53	32.69
Removal efficiency [2], %	-	90.96	90.88
CO_2 specific emission, kg/kWh	0.356	0.0356	0.0356

[1] composition (% by volume): Ar=0.94, CO_2=4.37, H_2O=3.36, N_2=78.82, O_2=12.52

[2] referred to the CO_2 generated by combustion.

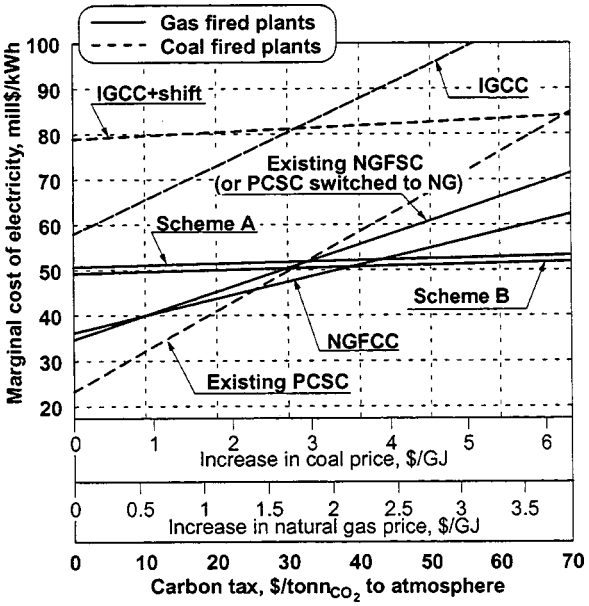


Fig. 3 Marginal cost of electricity (COE) versus carbon tax. For new plants yet to be built, the marginal COE incorporates all costs, including the cost of capital. For existing plants, marginal COE does not include amortization, whereby the marginal cost of existing NGFSCs is the same of switching existing PCSCs to natural gas.

Optimum Strategies to Reduce CO_2 Emissions From Power Plants

Given that conventional NGFCC are by themselves low CO_2 emitters (when compared to existing coal-fired or gas-fired steam plants), one might question whether the CCs with CO_2 removal considered here will ever constitute a meaningful proposition; in other words, whether realizing new plants with CO_2 removal like Scheme A or B is preferable to substituting existing plants with conventional and less troublesome NGFCCs. This issue, is best addressed by considering the following two scenarios:

1 new capacity must added under a constraint on overall CO_2 emissions from power stations
2 CO_2 emissions must be reduced while maintaining the same overall capacity

Under scenario (1), alternative technologies should be compared based on the marginal cost of the electricity generated by the new capacity. Under scenario (2), alternative technologies

Table 6 Breakdown of estimated investment costs

Plant configuration	NGFCC	Scheme A	Scheme B
Gas turbine, M$	54.6	54.6	54.6
Steam cycle, M$	92.1	81.5	88.5
CO_2 removal (DEA), M$		52.3	31.8
CO_2 compression, M$		5.6	5.5
Process facilities, M$	146.8	194.1	180.5
Balance of plant	12 % of process facilities		
Cost of engineering	18 % of process facilities		
Contingencies	5 %	10 % of process facilities	
Total cost, M$	198.2	271.7	252.7
Specific cost, $/kW	531	807	762

Table 7 Basic features and estimated cost of electricity (COE) of natural gas-fired and IGCC plants, with and without CO_2 sequestration. The fraction of COE attributable to the cost of fuel and CO_2 storage is estimated with reference to yearly averages of efficiency and specific emissions which differ by 5 percent from their design values; this 5 percent penalty is meant to account for ageing, fouling, and part-load operation.

Plant configuration		New plants CC	Scheme A	Scheme B	IGCC	IGGC + shift	Existing steam cycle plants	
Fuel		NG	NG	NG	Coal	Coal	NG	Coal
Net plant efficiency, %	design	56.11	50.61	51.09	45.95	39.29	40.00	39.00
	yearly average	53.31	48.08	48.54	43.65	37.33	38.00	37.05
Specific CO_2 emission, kg/kWh	design	0.356	0.0356	0.0356	0.709	0.0705	0.499	0.835
	yearly average	0.374	0.0374	0.0374	0.746	0.0742	0.525	0.879
Specific cost, $/kW		531	807	762	1536	1913	-	-
Interest during construction, %		3.92	8.05	8.05	16.97	16.97	-	-
COE* due to plant amortization, mill$/kWh		11.75	18.57	17.53	38.25	47.65	0.00	0.00
COE* due to O&M, mill$/kWh		2.11	3.56	3.47	5.71	6.99	3.36	6.75
COE* due to fuel, mill$/kWh		22.29	24.71	24.47	14.02	16.40	31.26	16.52
COE* due to storage, mill$/kWh		-	3.77	3.74	-	7.99	-	-
Total COE*, mill$/kWh		36.14	50.62	49.21	57.98	79.02	34.63	23.27

should be compared based on the cost of avoided CO_2 emissions. Given the scope of this paper, the discussion is limited to technologies based on fossil fuels.

(1) Addition of New Capacity Under a Constraint on CO_2 Emissions. For the sake of clarity, let's consider a capacity increase of 100 MW and let the constraint on CO_2 emissions be conceived as a limit of 0,0374 kg_{CO_2} per kWh generated in excess of the production of the existing system. This limit (one tenth the average specific emissions of NGFCCs: see Table 7) is so low that its implementation would virtually halt the increase of CO_2 emissions from power stations; on the other hand, its being equal to the average specific emissions of Scheme B allows easier comparisons without loosing generality. Under this constraint on the CO_2 emissions attributable to new capacity, one might conceive five basic strategies for the addition of the specified 100 MW:

(a) add 100 MW of new capacity with the characteristics of Scheme B
(b) add 100 MW of conventional NGFCCs without CO_2 removal and switch 95.1 MW of existing coal-fired steam plants to natural gas
(c) shut down 223.4 MW of existing gas-fired steam plants and add 323.4 MW of conventional NGFCCs without CO_2 removal
(d) shut down 66.7 MW of existing coal-fired steam plants and add 166.7 MW of conventional NGFCCs without CO_2 removal
(e) shut down 4.6 MW of existing coal-fired steam plants and add 104.6 MW of coal-fired IGCCs with shift and physical absorption for CO_2 removal

Due to the lower specific CO_2 emission of gas versus coal and of NGFCCs and IGCCs with shift versus the steam plants to be shut down (see Table 7), strategies (b) to (e) entail the same additional CO_2 emissions of strategy (a): for the assumed 100 MW, 3.75 $tonn_{CO_2}$ per hour.

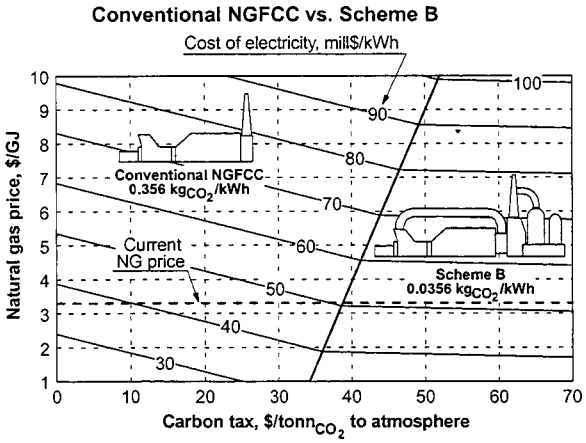


Fig. 4 Cost of electricity of conventional gas-fired CCs and Scheme B versus natural gas price and carbon tax. In the leftward region, where conventional CCs are more convenient, the lines at constant COE refer to NGFCCs; in the rightward region they refer to Scheme B.

Figure 5 shows the marginal cost of electricity[6] of the five strategies above versus the price of natural gas, for a price of coal of 1.7 and 1.0 $/GJ. The line for strategy (e) is horizontal because it corresponds to a configuration which requires no natural gas. Strategy (b)—adding conventional NGFCCs and switching coal-fired plants to NG to compensate for the additional CO_2 emissions

[6]Let B and B+E be the electricity production (kWh/yr) before and after the capacity increase, respectively; C_B and C_{E+B} the generation costs ($/yr) before and after the capacity increase, respectively. Then the marginal cost in Fig. 5 is the ratio between the extra cost ($C_{E+B}-C_B$) and the extra electricity E with respect to the base production B.

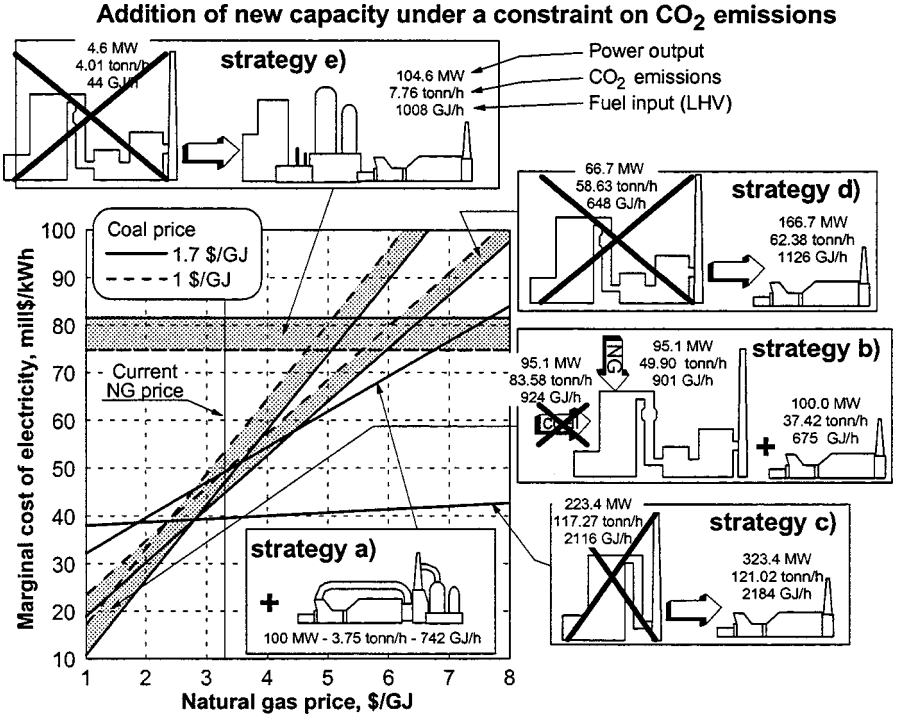


Fig. 5 Marginal COE of alternative strategies for the addition of 100 MW of new capacity under the constraint of virtually constant CO_2 emissions. All strategies (a) to (e) imply an increase of CO_2 emissions of 3.75 t/h, i.e., 0.0374 kg_{CO_2} per kWh generated in excess of the production of the existing system.

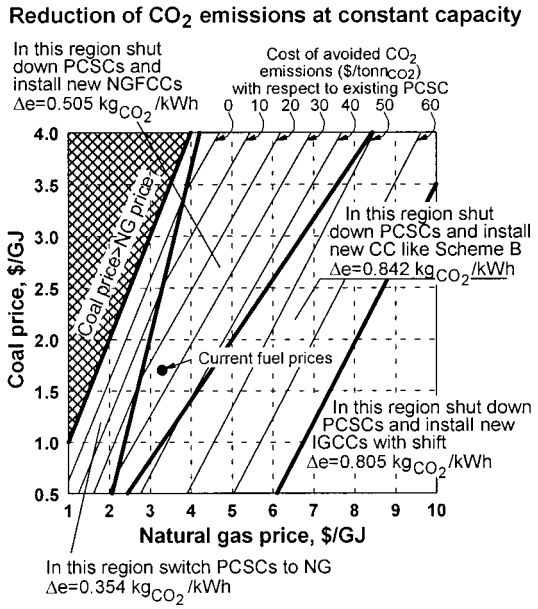


Fig. 6 Optimum strategies to reduce CO_2 emissions at constant generation capacity. Lines at constant cost of avoided CO_2 emission always refer to the cheapest strategy; in the rightmost region where IGCCs with shift and CO_2 removal should replace existing PCSCs, the cost of avoided CO_2 emissions is almost constant (about 69.2 $/tonn$_{CO_2}$). Δe is the reduction of specific emissions of each strategy.

of the new CC—is the strategy of choice for NG prices below about 3 $/GJ (not too far from current prices). For higher NG prices, the best strategy is always (c): shut down gas-fired steam cycles and install a new conventional NGFCC. In a system without gas-fired steam plants, the best strategy varies with the price of gas and (to some extent) of coal: for intermediate NG prices (between about 3 and 4 $/GJ, when coal price is 1.7 $/GJ), the best strategy is (d): add conventional NGFCCs and shut down PCSCs; for high NG prices (between about 4 and 7.5 $/GJ, when coal price is 1.7 $/GJ) the best strategy is (a): install a new CC like scheme B; for very high NG prices (above 7.5 $/GJ, when coal price is 1.7 $/GJ), the best strategy is (e): install a new IGCC with CO_2 removal.

At current fuel prices, CCs with CO_2 chemical absorption are third best behind shutting down NGFSCs and shutting down PCSCs; they would become second best at gas prices above 4.5 $/GJ or at coal prices around 1 $/GJ.

(2) Reduction of CO_2 Emissions at Constant Capacity. Under this scenario either existing plants are shut down to make room to lower CO_2-emission technologies or existing plants are fired with less carbon intensive fuels. Assuming that the existing system includes PCSC[7], it makes sense to consider the following four basic strategies:

(a) switch existing PCSCs to natural gas
(b) shut down existing PCSCs and install conventional NGFCCs without CO_2 removal
(c) shut down existing PCSCs and install CCs with CO_2 removal like Scheme B
(d) shut down existing PCSCs and install IGCCs with CO_2 removal

Figure 6 shows the merit of these four strategies as a function of fuel prices. The regions evidenced in the diagram mark the ranges of fuel prices for which one strategy is preferable to all others. The iso-cost lines running throughout the diagram give the cost of avoided CO_2 emissions, i.e., the increase of COE with respect to the reference PCSC (mills/kWh) divided by the reduction of specific emissions (kg_{CO_2}/kWh). For the sake of clarity, iso-cost lines report only the specific cost of the cheapest strategy; the cost of more expensive strategies is not reported.

For any given coal price, fuel switching (strategy (a)) is interesting only when the gas price is low (0.5 to 1.5 $/GJ more than coal). At higher gas prices the substitution of existing PCSCs becomes more appealing: first with conventional NGFCCs, then with plants like Scheme B and finally, at very high gas prices, with IGCCs equipped with CO_2 removal.

The cost per unit of avoided CO_2 emission shown on the diagram is also the credit (or the avoided carbon tax) one should get in order to make CO_2 removal economically viable. For example, at current fuel prices a credit of about 25 $/tonn of avoided CO_2 emission would drive toward the substitution of existing PCSCs with NGFCCs, which is coherent with the indications given by Fig. 3.

Conclusions

A natural gas fired CC can be coupled to a chemical absorption system to reduce specific CO_2 emissions by an order of magnitude and more. For systems based on large, ''FA'' heavy-duty turbines, CO_2 removal causes a reductions of net efficiency of about 5 percentage points, a reduction of power output by about 10 percent and an increase of the cost of electricity of about 13 mills/kWh. The best plant configuration includes recirculation of part of the flue gases to the gas turbine compressor.

CCs with CO_2 removal by chemical absorption become preferable to conventional NGFCCs for a carbon tax of $35–40 per tonn of CO_2 emitted to the atmosphere. A carbon tax between 25 to 35 $/tonn$_{CO_2}$ would drive toward shutting down existing PCSCs and install new conventional CCs; a tax below 25 $/tonn$_{CO_2}$ would be insufficient to stimulate the reduction of CO_2 emissions.

At current fuel prices, the most appealing strategy for the addition of new generation capacity under the constraint that CO_2 emissions from power generation are virtually constant is to shut down existing NGFSCs and install conventional NGFCCs. Second best is shutting down existing PCSCs and install conventional NGFCCs. The plants with CO_2 removal analyzed in the paper are third best; they would become second best at gas prices above 4.5 $/GJ or at coal prices around 1 $/GJ.

Acknowledgments

The authors are indebted to Dr. Giovanni Pelandini for his enthusiastic contribution in setting up the computer model of the power plant and obtaining the preliminary results.

Acronyms

CC = combined cycle
COE = cost of electricity
DEA = Diethanolamine
HHV = higher heating value
IGCC = integrated gasification CC
LHV = lower heating value
NG = natural gas
NGFCC = natural gas fired CC
NGFSC = natural gas fired steam cycle
PCSC = pulverized coal steam cycle

References

[1] U.S. Department of Energy, 1993, ''The Capture, Utilization and Disposal of Carbon Dioxide from Fossil Fuel-Fired Power Plants,'' report prepared for U.S. Dep. of Energy, contract No. DE-FG02-92ER30194.A000, July 1993.
[2] Kohl, A., and Riesenfeld, F., 1985, *Gas Purification*, 4th ed., Gulf Publishing, Houston.
[3] Chiesa, P., and Lozza, G., 1998, ''CO_2 Emission Abatement in IGCC Power

[7]If the existing system does not include any PCSC, the four options above can be re-stated by referring to oil-fired (or gas-fired) steam power plants.

Plants by Semiclosed Cycles. Part B—With Air-Blown Combustion and CO_2 Physical Absorption,'' ASME Paper 98-GT-385.
[4] Chiesa, P., and Lozza, G., 1998, ''CO_2 Emission Abatement in IGCC Power Plants by Semiclosed Cycles Part A—With Oxygen-Blown Combustion,'' ASME Paper 98-GT-384.
[5] Chiesa, P., Consonni, S., and Lozza, G., 1998, ''A Comparative Analysis of IGCCs with CO_2 Sequestration,'' *Proceedings Fourth Int. Conference on Greenhouse Gas Control Technologies*, Intertaken, Switzerland, Aug. 30–Sep. 2.
[6] Pelandini, G., and Michelutti, F., 1996, ''Thermodynamic Analysis of CO_2 Removal Systems from Combined Cycles Flue Gases,'' (in Italian) Graduation Thesis, School of Engineering of Politecnico di Milano, Milan, Italy, Academic Year 1994–95.
[7] Mathieu, P., and Nihart, R., 1998, ''Zero Emission Matiant Cycle,'' ASME Paper 98-GT-383.
[8] Audus, H., 1998, ''CO_2 Capture by Pre-Combustion Decarbonization of Natural Gas,'' *Proceedings Fourth Int. Conference on Greenhouse Gas Control Technologies*, Interlaken, Switzerland, Aug. 30–Sep. 2.
[9] Bannister, R. L., Newby, R. A., and Wen-Ching, Y., 1998, ''Final Report on the Development of a Hydrogen Fueled Combustion Turbine for Power Generation,'' ASME Paper 98-GT-201.
[10] Kizuka, N., et al., 1998, ''Conceptual Design of the Cooling System for 1700°C-Class Hydrogen-Fueled Combustion Gas Turbine,'' ASME Paper 98-GT-345.
[11] Lozza, G., 1990, ''Bottoming Steam Cycles for Combined Gas-Steam Power Plants: A Theoretical Estimation of Steam Turbine Performance and Cycle Analysis,'' *Proc. 1990 ASME Cogen-Turbo*, New Orleans, Louisiana, pp. 83–92.
[12] Consonni, S., 1992, ''Performance Prediction of Gas/Steam Cycles for Power Generation,'' Ph.D. thesis, No. 1893-T, Mechanical and Aerospace Engineering Dept., Princeton University, Princeton, NJ.
[13] Chiesa, P., Consonni, S., Lozza, G., and Macchi, E., 1993, ''Predicting the Ultimate Performance of Advanced Power Cycles Based on Very High Temperature Gas Turbine Engines,'' ASME Paper 93-GT-223.
[14] Macchi, E., Consonni, S., Lozza, G., and Chiesa, P., 1995, ''An Assessment of the Thermodynamic Performance of Mixed Gas-Steam Cycles: Part A—Intercooled and Steam-Injected Cycles,'' ASME J. Eng. Gas Turbines Power, **117**, pp. 489–498.
[15] ASPEN PLUS, 1996, release 9.3, reference manual.
[16] Miller, H. E., 1996, ''''F'' Technology—The First Half-Million Operating Hours,'' GE Publication GER-3950.
[17] Smelser, S. C., Stock, R. M., McCleary, G. J., 1991, ''Engineering and Economic Evaluation of CO_2 Removal from Fossil-Fuel-Fired Power Plant,'' EPRI Report IE-7365, Palo Alto, CA, USA.
[18] Hendriks, C., 1994, ''Carbon Dioxide Removal from Coal-Fired Power Plants,'' Ph.D. thesis, University of Utrecht, The Netherlands.
[19] Corti, A., Lombardi, L., and Manfrida, G., 1998, ''Absorption of CO_2 With Amines in a Semiclosed GT Cycle: Plant Performance and Operating Costs,'' ASME, Paper 98-GT-395.
[20] McMullan, J. T., et al., 1995, ''Techno-Economic Assessment Studies of Fossil Fuels and Fuel Woods Power Generation Technologies,'' *Joule II Program R&D in Clean Coal Technology*, report to the European Commission, Brussels, Belgium.
[21] *TAG—Technical Assessment Guide (Electric Supply)*, 1993, EPRI Report TR-102275 **1**, Rev. 7, Palo Alto (Ca), USA.
[22] Doctor, R. D., Molburg J. C., and Thimmapuram, P. R., 1996, ''KRW Oxygen-Blown Gasification Combined Cycle: Carbon Dioxide Recovery, Transport and Disposal,'' DOE report prepared by Energy System Division Argonne National Laboratory, August 1996.
[23] Chiesa, P., and Consonni, S., 1998, ''Shift Reactors and Physical Absorption for low-CO_2 Emission IGCCs,'' ASME Paper 98-GT-396.
[24] Fujoka, Y., et al., 1997, ''Cost Comparison in Various CO_2 Ocean Disposal Options,'' Energy Convers. Manage., **38**, pp. S273–S277.

Handling of a Semiclosed Cycle Gas Turbine With a Carbon Dioxide-Argon Working Fluid

Inaki Ulizar
Industria de Turbopropulsores—Ajalvir,
Torrejon de Ardoz,
Madrid, Spain

Pericles Pilidis
School of Mechanical Engineering,
Cranfield University,
Cranfield, Bedford, U.K.

This paper outlines the handling of a semi-closed cycle gas turbine, its working fluid is carbon dioxide and the fuel is low heating value gas from coal; however, at startup air and natural gas are used. The objective of the machine is to produce clean electricity with the smallest efficiency penalty. Many aspects of the operation of the engine are examined in this paper; these include starting requirements, stator vane and bleed valve scheduling, and the working fluid transition from air to carbon dioxide. Other features highlighted are the compressor operating lines and surge margins. The present paper describes the salient features of the three main stages into which the engine operation has been divided. These stages are: startup to synchronous idle, change of working fluid (from air to Carbon Dioxide-Argon) and fuel (from natural gas to coal synthetic gas) at synchronous idle, and part load operation. Preliminary findings show that engine handling can be carried out effectively with variable stators. This is possible because of the two-shaft gas generator. Another point of interest is the large increase of corrected speed relative to rotational speed experienced when the working fluid changes from air to carbon dioxide. In general the control of the engine does not seem to present any insurmountable problems despite the complexities arising from the need to change working fluid and fuel. [S0742-4795(00)02903-3]

Introduction

The continuing concern over the emission of greenhouse gases coupled with the ever-increasing demand for electrical energy poses a very difficult challenge to power engineers. One possible solution to these conflicting requirements, if solid fossil fuels are to be employed, is to collect and dispose of the emissions in a controlled way.

For example, an internal combustion semi-closed power cycle, where the working fluid is carbon dioxide, with oxygen and fuel injected in the combustion chamber and excess carbon dioxide collected at the outlet seems to be a very attractive environmental proposition. Such a cycle has a dual advantage, collecting, and storing safely the excess carbon dioxide, the emission of greenhouse gases is controlled. Having a clean fuel and no air in the working fluid, there are no emissions of oxides of sulphur and nitrogen. This proposition becomes particularly attractive when coal is considered. This fuel is in plentiful supply and when gasified it can be a very clean source of energy. The major drawbacks of such a cycle are the complexity of the equipment and the significant reduction in efficiency caused by the need to produce pure oxygen and to produce and use low heating value fuels.

The selection of the main-cycle details, a preliminary assessment of efficiencies and sizing of the gas turbine have been documented in [1]; where also the selection of thermodynamic parameters for efficiency and specific power can be found. In that paper, the basis of the design was a two spool gas turbine, consisting of Low and High Pressure Compressors (LPC and HPC) and a High Pressure and a Low Pressure Turbine (HPT and LPT), the latter also being the Power Turbine. The engine layout is shown in Fig. 1. The focus of the present paper is the detailed off-design performance of the engine, from starting, through the change of fuel and working fluid to part load performance.

The analysis was carried out using GTSI [2] an advanced gas turbine simulation program working on the basic principles of TURBOMATCH [3], the Cranfield gas turbine performance simulation software. GTSI incorporates the ability to use a variety of engine working fluids and fuels. It incorporates the properties of many gases and their mixtures [4] and can be used for a variety of tasks [5,6,1].

The Engine

This paper describes an internal combustion turbine where the working fluid is carbon dioxide, the main mass of which is recirculated in the engine. The fuel based on coal, and the requirement to make carbon dioxide collection an easy process dictates the choice of working fluid.

Heat addition is achieved by injecting oxygen and coal gas fuel in the combustion chamber. To keep a constant mass flow through the device, water is collected at the outlet of the cooler and some carbon dioxide is extracted at the HPC delivery. The gas extracted is compressed to at least 60 atmospheres for efficient collection. All calculations include this compression work, the fuel compression work and the energy drain from the air separation plant that provides the oxygen.

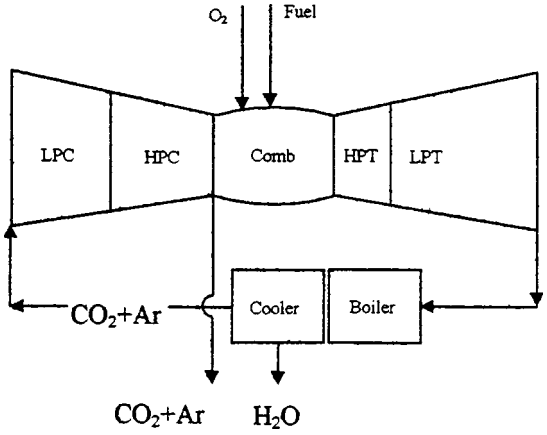


Fig. 1 The semi-closed cycle engine

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-374. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

Table 1 Engine features

LPC Press. Ratio	7
HPC Press. Ratio	8
Overall Press. Ratio	56
HPT Press. Ratio	3.2
LPT Press. Ratio	15.3
LPC Stages	9
HPC stages	12
HPT stages	2
LPT Stages	6
LPC inlet Temp (K)	300
HPC exit Temp (K)	708
HPT inlet Temp (K)	1650
LPT exit Temp (K)	960
Gas Turbine Power (MW)	161
Steam Turbine Power (MW)	103
Nett Output (MW)	193
Mass Flow (kg/s)	500
Fuel Flow (kg/s)	52
ηgt	0.172
ηcc	0.367
LPC inlet diameter (m)	1.75
LPT outlet diameter (m)	2.14
Overall gas turbine length (m)	5.4
RPM HPC	5000
RPM LPC	3000

Figure 1 shows the engine analyzed in this paper. The selection of the main cycle details is documented in [1]. A brief summary is given in Table 1. Three main aspects of the handling of the engine are investigated:

1 starting to synchronous idle
2 change of working fluid at idle
3 part load performance

Compressor Running Lines

The compressor characteristics throughout the operating range are shown for all three phases in Fig. 2. Figure 2(*a*) shows the LPC compressor characteristics and operating line. Two LP compressor maps are shown, that (with the high values of mass flow) corresponds to the fully open LPC stators, while the other corresponds to the closed stators. The intermediate maps resulting from intermediate positions of the variable compressor stators have been omitted for clarity. The change in LPC stators is shown wherever it is appropriate.

The same comments can be made for the HP compressor, the paths and characteristics shown in Fig. 2(*b*). The compressor maps shown can be applied to any situation because the nondimensional parameters contain the appropriate gas constants to cater for a change of fluid. The stators of both compressors are actuated so that the engine has adequate surge margin throughout the operating range. In the LPC they also satisfy a significant requirement of mass flow modulation. The present predictions seem to indicate that bleed valves are not required. However, this needs to be ascertained with more detailed investigations. Three critical points have been identified in the operating regime. They are, two synchronous idle points, one with air (point 1 in Figs. 2(*a*) and 2(*b*), the other with carbon dioxide (point 2) and the design point (point 3).

The LPC running line, shown in Fig. 2(*a*), needs extensive stator actuation to achieve the large mass flow modulation required. The HPC (Fig. 2(*b*)) exhibits a more conventional running line. This means that the variable geometry actuation requirements for the stators of the HPC will be much smaller than for those of the LPC, which needs to accommodate a large mass flow increase at constant speed (points 2 to 3 in Fig. 2(*a*)). It can be noticed that before point 1 is reached, the HPC trajectory is such that mass flow rises, then falls and finally it starts rising again. This can also be observed in Fig. 3(*b*), where rotational speed is shown. This occurs while the LPC is being ''pulled'' by the HP system. A more traditional trajectory could be achieved by a better optimization of the stator schedules. In the present case, the requirement was to ensure a surge margin of at least 10 percent with the characteristics employed.

During the phase of change of working fluid (point 1 to 2 in Figs. 2(*a*) and 2(*b*)) there is a large change of the LPC corrected flow. This is mainly due to the increase in mass flow and partly due to the change of properties. The HPC stators remain in the same setting after the synchronous idle setting with air is reached. Changes in corrected flow are achieved by changing rotational speed.

During the phase of change of working fluid (point 1 to 2 in Figs. 2(*a*) and 2(*b*)) the change of the HPC corrected flow is much smaller than that of the LPC. This is because the large change in mass flow is compensated by the large increase of the LPC pressure ratio, which increases from approximately 2.5 to nearly 6 (this can be observed in Fig. 2(*a*)).

Starting

Starting is considered to be the phase beginning with the engine stationary and ending with the engine at synchronous speed and no load, with air as working fluid and natural gas fuel, so in this case the oxidizer is contained in the air stream. The engine is started from rest by external power, coupled to the HP shaft, and brought to synchronous speed with natural gas fuel and air.

This phase starts with the rotor at rest. Power is provided from an external source and brings the engine to a self sustaining condition. The engine is then brought to an operating point where the LP shaft speed is 3000 RPM.

The performance of the engine in these conditions, as with all gas turbines, is very difficult to predict because the performance of the components is not well understood.

Using the predictions of this and previous work, the authors estimate that a starting device of about 1 MW will be required to bring the LP shaft to a speed of about 70 percent. This speed and power was chosen after examining Fig. 3(*a*). It appears to be at this point that there is an acceptable level of performance of the components.

HP shaft speed in these conditions is expected to be about 75 percent (Fig. 3(*b*)). Figure 3(*c*) shows the variable stator setting of both compressors, during the starting phase. Figure 3(*d*) shows the turbine temperatures in these conditions. As expected, relatively high temperatures are encountered during this phase. This is resulting from the relatively large amount of heat input required to overcome the large component losses arising from operating at conditions so different to those of the design point. Finally, Fig. 3(*e*) shows the fuel flow required for this phase. The LP shaft speed has now reached 3000 RPM and this rotational speed will not change any more, because the LP turbine is synchronized with the electricity generator.

Change of Working Fluid at Idle

Once the engine has stabilized at synchronous idle, the working fluid is changed from air to carbon dioxide. The main reason for doing this now is that at the end of this phase the oxidizer will be oxygen, fed through the main oxygen injectors within the engine combustors. If the change of working fluid takes place at synchronous idle, it is expected that the capacity of the oxygen injectors will not be exceeded, even if hyperventilation is required.

The transient process of changing the working fluid was investigated as a sequence of quasi steady state points. The time dimension is not considered here. A detailed transient study must be carried out later when it is necessary to understand the dynamic processes governing this phase. However, the present analysis gives a good picture of the handling needs of the engine.

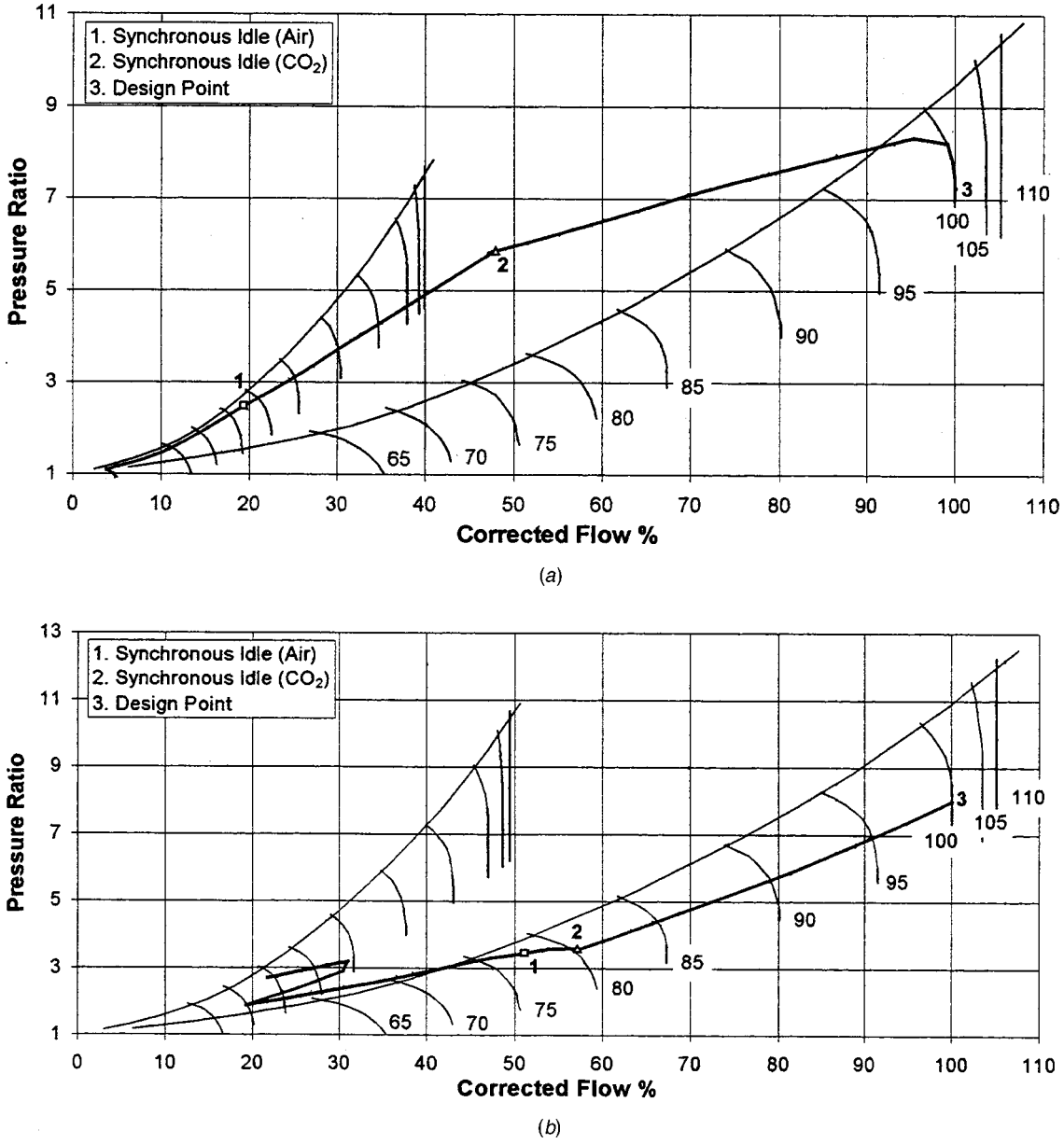


Fig. 2 (*a*) Off design performance—LPC operating line; (*b*) off design performance—HPC operating line

All the engine parameters change, except the LP shaft speed, which is 3000 RPM. This is, naturally, the matching criterion. In addition a minimum level of surge margin was a requirement.

It can be seen in Fig. 4(*a*) that the HPC rotational speed falls from about 96 to 81 percent. In this figure it is also interesting to observe the large increase of corrected speed relative to rotational speed. This is due to the large difference of the properties of air and carbon dioxide. It can be seen that LPC corrected speed increases by about 20 percent while rotational speed is constant. In the HPC the rotational speed falls by about 15 percent while corrected speed rises. This finding indicates that the change of working fluid should be carried out at idle to avoid rotor overdesign to allow for overspeeds at full power.

The turbine temperatures fall slightly as shown in Fig. 4(*b*). The mass flow (Fig. 4(*c*)) and the fuel flow (Fig. 4(*d*)) both rise. The LPC variable stator is opened slightly, while the HPC stators remain at a constant setting.

At the end of this phase, the fuel will be switched from natural gas to synthetic gas. In these circumstances it may be necessary to have a rich oxygen mixture to ensure combustion of the fuel while the change is taking place. The difference in performance is relatively small. The small changes can be detected when comparing the end points of the phase where the working fluid is carbon dioxide (Figs. 4) and the part load performance of the engine. Compressor stall margin includes a margin for the reduction caused by the larger difference between compressor and turbine mass flow resulting from the use of low calorific value gas. As stated above the engine control laws employed required a minimum surge margin of 10 percent in both compressors.

If the compressor trajectories are examined, it can be observed that the LPC sees a large change in mass flow (Fig. 2(*a*), points 1 and 2). In the case of the HPC, the change in corrected mass flow is small (Fig. 2(*b*)) because the LPC pressure ratio change is very large. In the HP spool it is interesting to see that although shaft speed falls significantly (Fig. 4(*a*)), the corrected speed rises (Fig. 2(*b*)). This is due to the change of gas properties that affect the corrected speed. In the case of the LPC the same was observed, but the change is not evident in Fig. 2(*a*) because the stator angle was changed. However, in both cases there is a notable increase of corrected shaft speed relative to physical shaft speed.

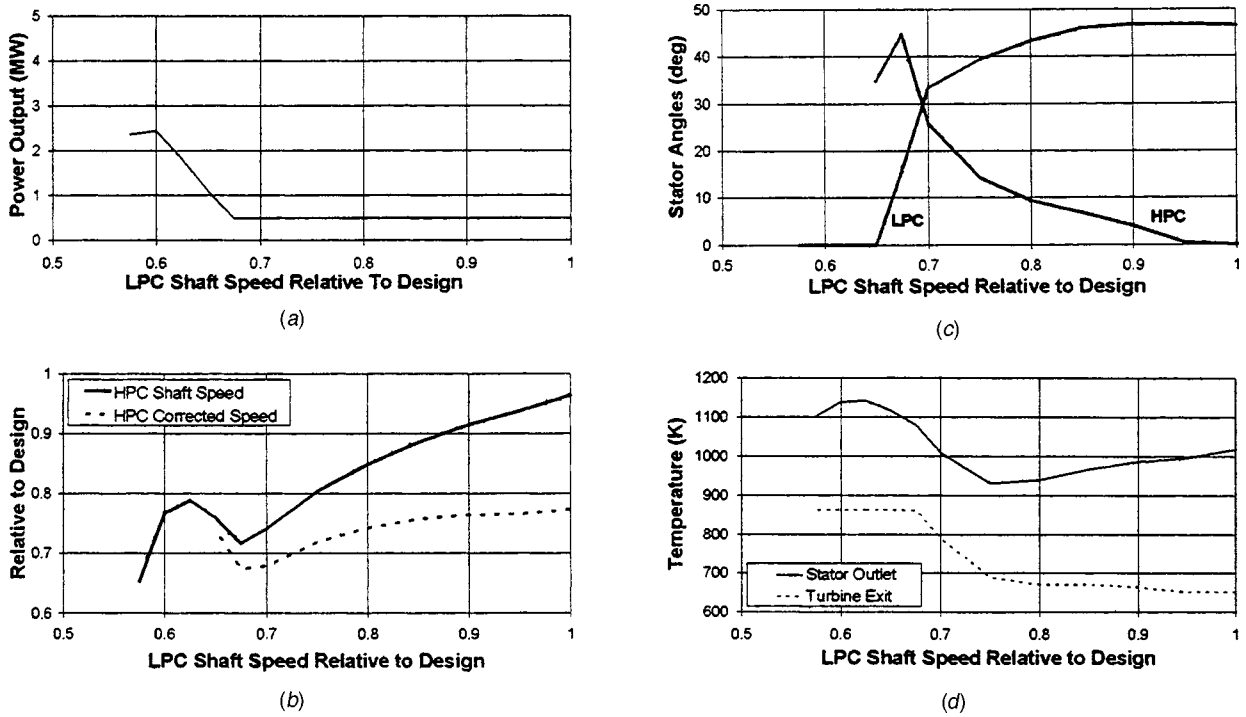


Fig. 3 (*a*) **Gas turbine power output, starting;** (*b*) **starting—HPC shaft speed;** (*c*) **starting—variable stator angles; and** (*d*) **starting temperatures**

Part Load Performance

Once the engine has been synchronized it can produce useful power that will be absorbed by the electricity generator. The power output will vary up to the design point power output (point 3 in Figs. 2 and Fig. 5(*a*)) as the turbine temperatures rise nearly linearly with HP spool speed to their design values (Table 1). Figure 5(*a*) shows the output of the semiclosed cycle (labeled Gas Turbine) and of the same plant with a steam bottorning cycle (labeled Combined Cycle).

The LPC stators open (Fig. 5(*b*)) to allow a progressively larger

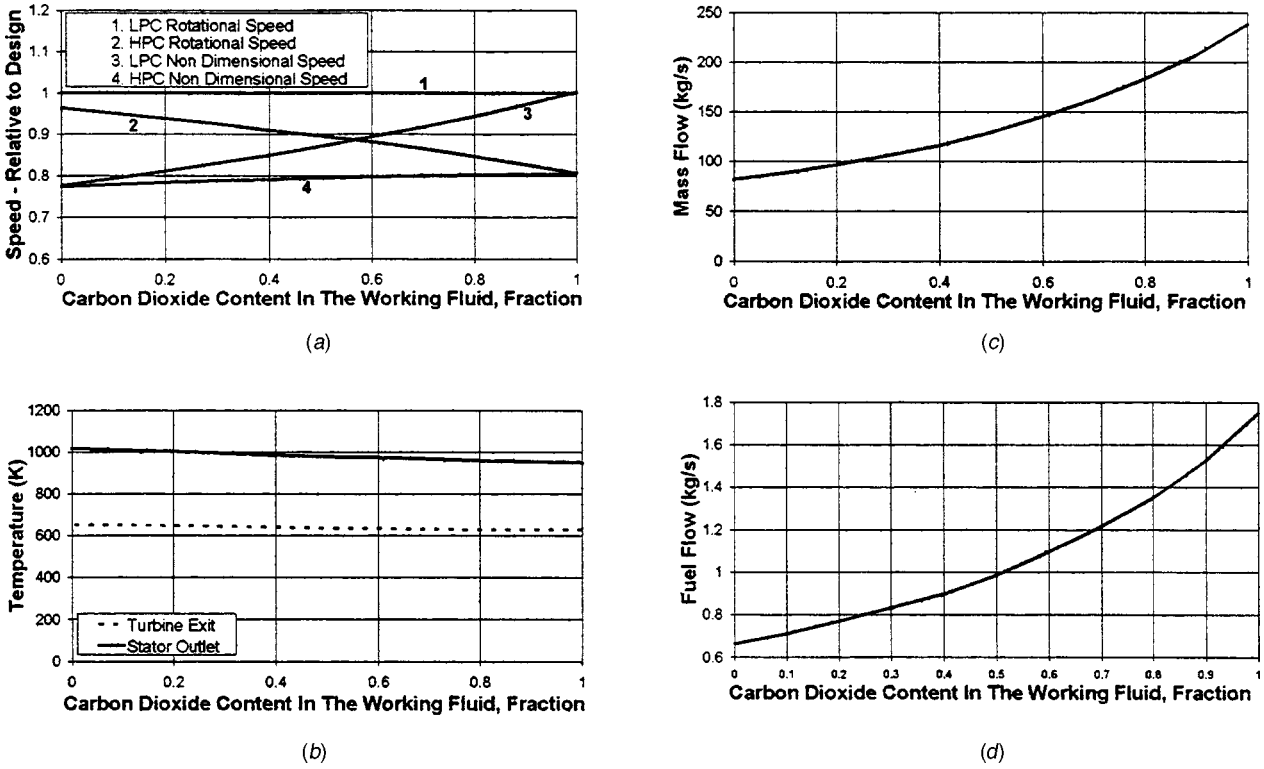
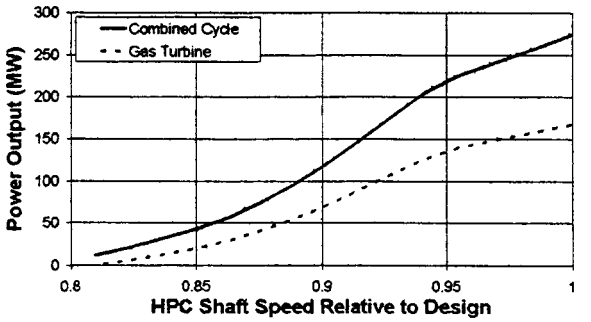
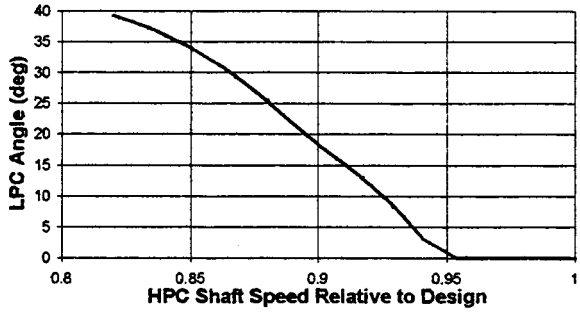


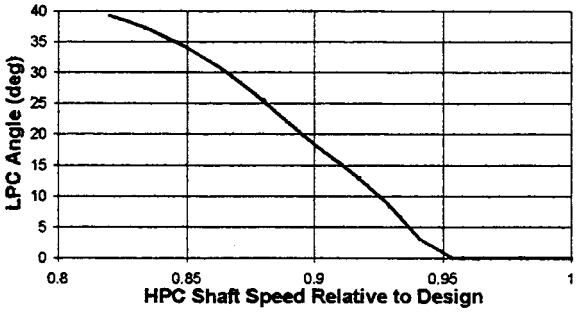
Fig. 4 (*a*) **Change of working fluid—change of speeds;** (*b*) **change of working fluid—temperatures;** (*c*) **change of working fluid—mass flow; and** (*d*) **change of working fluid—fuel flow (N. Gas)**



(*a*)



(*b*)



(*c*)

Fig. 5 (*a*) Part load performance—power output; (*b*) part load performance—LPC stator angle; and (*c*) part load performance—mass flow

mass flow (Fig. 5(*c*)) to enter the engine. This change of mass flow is required to match properly the two compressors. The HPC stators are at their fully open setting from the phase where the working fluid was changed to carbon dioxide and they require no further actuation. This rotor is not coupled to the generator, so its rotational speed can change while the engine is synchronized.

It is expected that all these changes will take place slowly, as in all large gas turbines. The large thermal mass of the engine will need a considerable heat soakage period.

Figure 2(*a*) shows the large change in mass flow (point 2 to point 3) achieved in the LPC, at constant speed, by changing the stator setting (shown in Fig. 5(*b*)). The HPC exhibits a more traditional working line (Fig. 2(*b*)).

Conclusion

In the present paper, the prediction of the performance of a simple cycle semi-closed cycle gas turbine has been outlined. Three main phases in the off-design operation of the engine are described in detail, these are

1 starting to synchronous idle
2 change of working fluid at idle
3 part load performance

It appears that with suitable actuation of the compressor stators, this two spool gas turbine may be taken throughout the complete operating range. Due to the uncertainty of the performance of the components at very low speeds, however, it is not possible to discount the need for bleed valves at this stage.

A noteworthy conclusion is that observed in the phase where the working fluid is changed from air to carbon dioxide. There is a large increase of corrected speed relative to speed. This indicates that to avoid the structural overdesign of the rotors, the change of working fluids should take place at synchronous idle.

The next stage of the work is to analyze the detailed performance of the components, specially the compressors. The analysis of the components could be executed using techniques such as those described in [7,8].

The present exercise reveals that the control of such an engine does not appear to present any challenges beyond those of current, state of the art (in 1998) equipment.

Acknowledgment

The authors are grateful to the C.E.C. for the financial support for the initial phase of this task and to their colleagues in I.T.P. and Cranfield for their support and encouragement. The assistance of Mr. Kenneth Sargent in the production of this paper is also acknowledged.

References

[1] Ulizar, I., and Pilidis, P., 1997, ''Design of a Semiclosed Cycle Gas Turbine with Carbon Dioxide-Argon as Working Fluid,'' ASME Paper 97-GT-125.
[2] Ulizar, I., and Pilidis, P., 1996, ''GTSI, Simulation of Gas Turbines with Unorthodox Working Fluids and Fuels,'' *Turbomacchine 96*. Italian Thermotechnical Conference, Genova.
[3] Palmer, J. R., 1983, ''The TURBOMATCH scheme for Gas Turbine Performance Calculations,'' users' guide, Cranfield Institute of Technology, United Kingdom.
[4] Keenan, J. H., Kaye, J., and Chao, J., *Gas Tables*, International Version.
[5] Hunter, I. H., 1994, ''Design of Turbomachinery for Closed and Semiclosed Gas Turbine Cycles,'' M.Sc. thesis, Cranfield University, United Kingdom.
[6] Navaratnam, M., 1994, ''The Investigation of an Aeroderivative Gas Turbine Using Alternative Working Fluids in Closed and Semiclosed Cycles,'' M.Sc. thesis, Cranfield University, United Kingdom.
[7] Massardo, A., and Satta, A., 1990, ''Axial Flow Compressor Design Optimization: Part. I—Pitchline Analysis and Multivariable Objective Function Influence,'' ASME J. Turbomach., **112**, pp. 399–404.
[8] Massardo, A., Satta, A., and Marini, M., 1990, ''Axial Flow Compressor Design Optimization: Part II—Through Flow Analysis,'' ASME J. Turbomach., **112**, pp. 405–410.

Performance of Pre-Swirl Rotating-Disc Systems

Hasan Karabay
Robert Pilbrow[1]
Michael Wilson
J. Michael Owen

Department of Mechanical Engineering,
Faculty of Engineering and Design,
University of Bath,
Bath BA2 7AY, UK

This paper summarizes and extends recent theoretical, computational, and experimental research into the fluid mechanics, thermodynamics, and heat transfer characteristics of the so-called cover-plate pre-swirl system. Experiments were carried out in a purpose-built rotating-disc rig, and the Reynolds-averaged Navier-Stokes equations were solved using two-dimensional (axisymmetric) and three-dimensional computational codes, both of which incorporated low-Reynolds-number k-ε turbulence models. The free-vortex flow, which occurs inside the rotating cavity between the disc and cover-plate, is controlled principally by the pre-swirl ratio, β_p: this is the ratio of the tangential velocity of the air leaving the nozzles to that of the rotating disc. Computed values of the tangential velocity are in good agreement with measurements, and computed distributions of pressure are in close agreement with those predicted by a one-dimensional theoretical model. It is shown theoretically and computationally that there is a critical pre-swirl ratio, $\beta_{p,crit}$, for which the frictional moment on the rotating discs is zero, and there is an optimal pre-swirl ratio, $\beta_{p,opt}$, where the average Nusselt number is a minimum. Computations show that, for $\beta_p < \beta_{p,opt}$, the temperature of the blade-cooling air decreases as β_p increases; for $\beta_p > \beta_{p,opt}$, whether the temperature of the cooling air increases or decreases as β_p increases depends on the flow conditions and on the temperature difference between the disc and the air. Owing to the three-dimensional flow and heat transfer near the blade-cooling holes, and to unquantifiable uncertainties in the experimental measurements, there were significant differences between the computed and measured temperatures of the blade-cooling air. In the main, the three-dimensional computations produced smaller differences than the two-dimensional computations. [S0742-4795(00)01902-5]

1 Introduction

Pre-swirl nozzles are used in many gas turbines to supply the cooling air to the turbine blades. The stationary nozzles swirl the air in the direction of rotation of the turbine disc, thereby reducing the temperature of the air relative to the blades. In the so-called cover-plate pre-swirl system shown in Fig. 1, the pre-swirl nozzles are located at a radius $r=r_1$ and the entrance to the blade-cooling passages is at $r=r_2$, where $r_2>r_1$. The swirling air flows radially outwards in the rotating cavity between the rotor (or rotating disc) and a cover-plate attached to it.

Pre-swirl systems have been studied by many research workers, and the reader is referred to the work of Meierhofer and Franklin [1], El-Oun and Owen [2], Chen et al. [3,4], Popp et al. [5], Wilson et al. [6], Karabay et al. [7], and Pilbrow et al. [8]. Some of the authors' recent results for cover-plate pre-swirl systems are briefly reviewed below.

Karabay et al. [7] showed, computationally and experimentally, that the flow between the rotating cover-plate and the stator was like a conventional rotor-stator system, in which a disc rotates close to a stationary casing (see Owen and Rogers [9]), and the flow between the cover-plate and the rotor was like that in a rotating cavity in which two discs corotate (see Owen and Rogers [10]). In particular, at the high coolant flow rates that occur in practical cases, free-vortex flow occurs outside the boundary layers in the rotating cavity. Pilbrow et al. [8] extended this work to include heat transfer from the rotor to the cooling air, and computations of the local Nusselt numbers were in reasonably good agreement with measured values for a wide range of pre-swirl ratios, β_p, nondimensional coolant flow rates, C_w, and rotational Reynolds numbers, Re_ϕ.

Karabay et al. [11] conducted a combined theoretical and computational study of the flow and heat transfer in the ''simple ro-

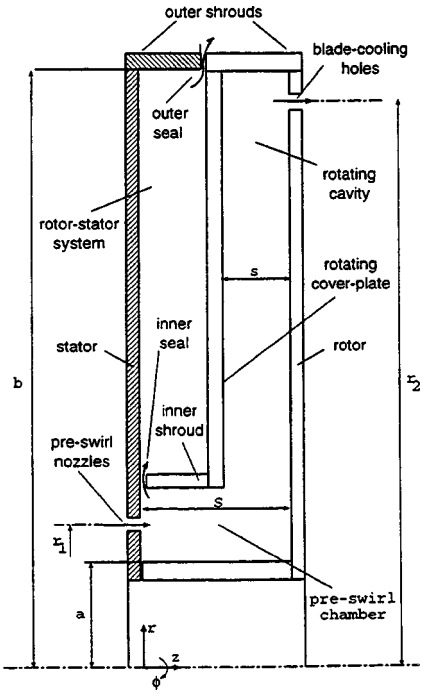


Fig. 1 Simplified diagram of cover-plate pre-swirl system (for the rig used in this study: $a=80$ mm, $b=207$ mm, $r_1=90$ mm, $r_2=200$ mm, $S=25$ mm, $s=10$ mm)

[1]Now at Rolls Royce plc, Bristol, UK.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-197. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters January 3, 2000. Associate Technical Editor: D. Wisler.

tating cavity,'' disregarding the rest of the cover-plate system. They derived an expression for the radial distribution of pressure, and used the Reynolds analogy to derive a relationship for the adiabatic-disc temperature. For the case where there is free-vortex flow throughout the rotating cavity, there is a critical pre-swirl ratio, $\beta_{p,\mathrm{crit}}$, for which the moment coefficient, C_m, is zero, and an optimal pre-swirl ratio, $\beta_{p,\mathrm{opt}}$, for which the average Nusselt number of the heated rotor is a minimum. Karabay's [12] thesis is concerned with the computation of flow and heat transfer in pre-swirl systems, and this paper presents some of these computations.

It is the object of this paper to extend the previous work and to provide some guidance for the designers of internal cooling-air systems. In particular, the designer is interested in the pressure drop associated with pre-swirl systems and needs to know the effect of β_p on the temperature of the blade-cooling air when heat transfer from the turbine disc is significant. Previous computational work has concentrated on axisymmetric systems, but three-dimensional effects can be significant near the blade-cooling holes in the rotor, as discussed below.

Section 2 of the paper outlines the computational method and experimental apparatus used in this study. The fluid dynamics of the cover-plate system shown in Fig. 1 is discussed in Section 3, and the thermodynamics and heat transfer in Section 4. The conclusions are summarized in Section 5.

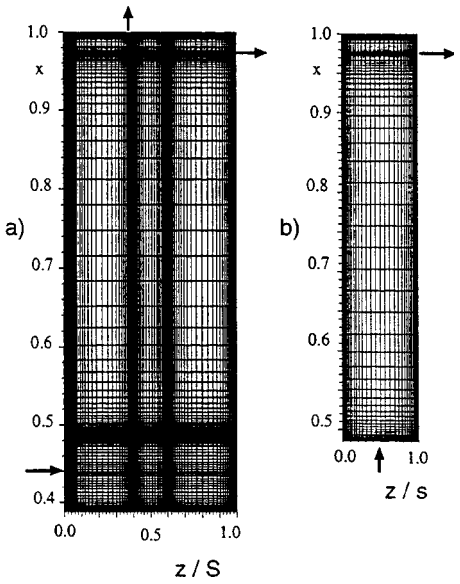


Fig. 2 Grid distributions used for computations: (*a*) whole system; (*b*) simple cavity.

2 Computational and Experimental Methods

2.1 Computational Model. The computational model has been described by Karabay et al. [7] and further details are given by Karabay [12]. For completeness, the salient points of the model are given below.

Turbulent flow computations were made using the low-Reynolds, number k-ε models of Launder-Sharma [13] and of Morse [14]. Both used gradient-diffusion modeling of transport, and the latter model incorporated modifications suggested by its author (see Chen et al. [15] for details). Incompressible flow was assumed and turbulent heat transfer was represented using a turbulent Prandtl number Pr_t equal to 0.9.

A staggered grid was used in which the axial and radial velocity components were stored midway between the grid-points where other solution variables were located (pressure, tangential velocity, turbulence kinetic energy and dissipation rate, and total enthalpy). The cover-plate and inner shroud (see Fig. 1) were represented by block obstructions within the computational grid and the equations were solved using the SIMPLEC pressure-correction algorithm.

The inlet nozzles and blade-cooling holes of the experimental rig were represented in the axisymmetric model by equivalent-area annular slots on the stator and rotor, with centerlines at $r=r_1$ and $r=r_2$, respectively. The axial velocity of the pre-swirl air, $V_{z,p}$, which was assumed to be uniform at the inlet, was deduced from the prescribed flow rate. Similarly, the axial velocity at the blade-cooling slot and the radial velocity at the outer seal were calculated from prescribed flow rates.

The inlet tangential velocity, $V_{\phi,p}$, was fixed to give the required pre-swirl ratio β_p, and Neumann (zero normal-derivative) boundary conditions for tangential velocity were used at the two outlets ($\partial V_\phi/\partial z=0$ at the blade-cooling slot and $\partial V_\phi/\partial r=0$ at the outer seal). The remaining velocity components at flow boundaries were taken to be zero, and no-slip conditions were applied at all solid surfaces. The inlet temperature of the pre-swirl air and the surface temperature distribution of the rotor (see Fig. 1) were taken from measured values (Karabay et al. [7]). The surfaces of the shrouds, the stator, the cover-plate and the inner casing at $r=a$ were all assumed adiabatic. At the blade-cooling slot, $\partial T/\partial z=0$ was assumed, and $\partial T/\partial r=0$ was used for the flow leaving at the outer seal.

The turbulence models required a very fine grid near the boundaries, with $y^+<0.5$ for the near-wall grid nodes, and the grid-spacing increased geometrically away from walls (including the cover-plate and shroud) with expansion factors of about 1.2. Computations were conducted for the ''whole system,'' corresponding to the complete cover-plate system shown in Fig. 1, and for the smaller ''simple cavity,'' corresponding to the rotating cavity between the rotor and the cover-plate shown in Fig. 1. Only two-dimensional (axisymmetric) computations were carried out in the ''whole system'' the ''simple cavity'' was used for the two-dimensional and three-dimensional computations described in Section 4.4. A 223 by 223 axial by radial grid was used for the ''whole system''; for the ''simple cavity'' a 67 by 111 axial by radial grid was used (see Fig. 2). (Computations were also conducted with a 141×185 grid, which made no significant difference to the comparison between the computed and measured Nusselt numbers; this suggests that the results presented below are sensibly grid-independent.)

2.2 Experimental Apparatus. A diagram of the rotating-disc rig is shown in Fig. 3, and details of the apparatus are given by Pilbrow et al. [8]. The main features of the apparatus are summarized below, and the total-temperature probes are described in detail. Although two pressure transducers were fitted to the cover-plate, an accident with the rig damaged the wiring and so it was not possible to obtain pressure measurements.

2.2.1 Rotating-Disc Rig. The outer radius of the system, b, was approximately 207 mm, and the radial locations of the pre-swirl nozzles and blade-cooling holes were $r_1=90$ mm and $r_2=200$ mm. There were 19 pre-swirl nozzles of 7.92 mm diameter, angled at 20 deg to the tangential direction, and 60 blade-cooling holes of 7.7 mm diameter, with their axes normal to the rotor. The pre-swirl nozzles were made from stainless-steel tubes, which were bent through an angle of 70 deg to produce the required swirl angle. The outlet ends of the tube were ground so that they were flush with the surface of the stator. The outlet holes from the pre-swirl nozzles were therefore elliptic when viewed from the axial direction. The axial spacing between the 5-mm-thick cover-plate and the stator was 10 mm, and between the cover-plate and the rotor the spacing was also 10 mm.

The rotor was of composite construction, comprising a steel front disc, two layers of Rohacell foam and an aluminum back disc. A 1-mm-thick fibre-glass mat was bonded to the front (cavity-side) face of the steel disc, which was 10-mm thick, and 10 thermocouples and 6 fluxmeters were embedded in the surface

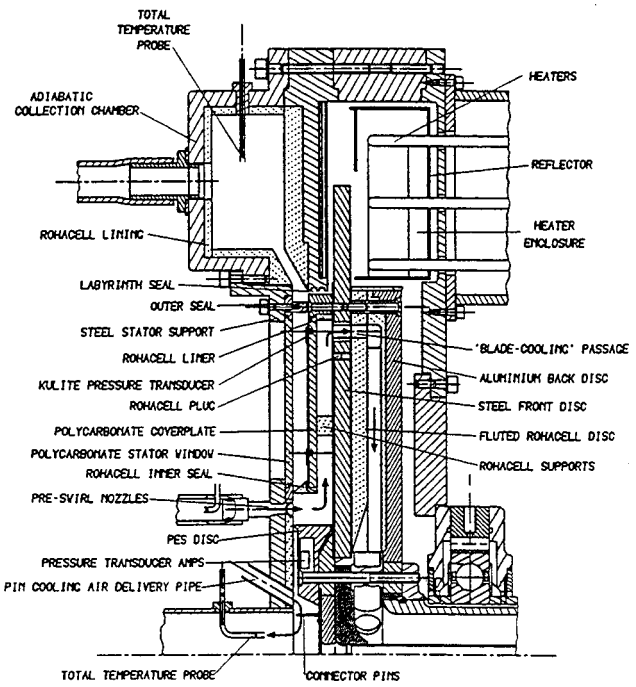


Fig. 3 Diagram of the rotating-disc rig

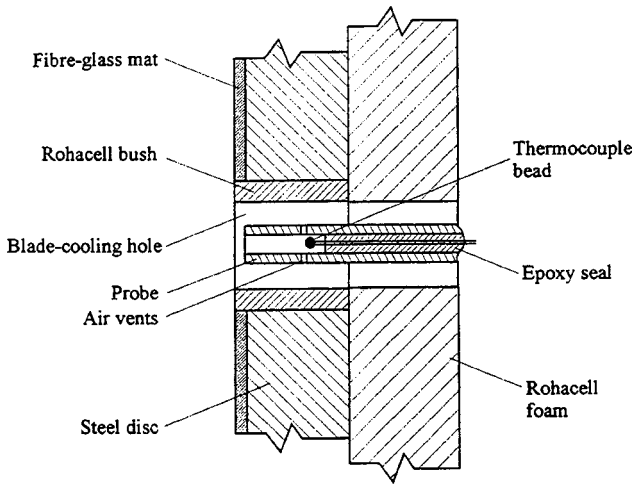


Fig. 4 Details of total-temperature probe located in blade-cooling hole (not to scale)

of the mat. The Rohacell foam provided thermal insulation, for the back of the steel disc, and contained passages for the blade-cooling air; each layer of foam was approximately 10-mm thick. The signals from all the rotating instrumentation were brought out through a 24-channel silver/silver graphite slipring unit, and the voltages were measured with a precision of ± 1 μV by a computer-controlled Solartron data logger.

The outer part of the rotating disc was heated up to 150°C by means of stationary radiant electric heater units, with a maximum power output of 9.5 kW. The actual temperature distribution over the instrumented section of the disc depended on the heat conducted radially inwards from the periphery. The maximum temperature in the instrumented section of the disc was around 70°C, but it depended on the rotational speed and on the flow rate of the cooling air. The temperature of the air at the pre-swirl nozzles was maintained at between 10°C and 20°C.

For a typical engine, $\mathrm{Re}_\phi \approx 10^7$, $\lambda_T \approx 0.3$, $\beta_p \approx 2.5$. For the heat transfer tests in the rig, the following range of parameters was tested: $0.5\times 10^5 < \mathrm{Re}_\phi < 1.6\times 10^6$, $0.16 < \lambda_T < 0.32$, $0.9 < \beta_p < 3.1$. It is shown in Section 3.1 that the flow structure depends principally on λ_T and β_p. The rig is therefore able to model the flow structures expected in engines, although the Nusselt numbers and temperatures measured in the rig are much smaller than those found in engines.

2.2.2 Total-Temperature Probes. The blade-cooling holes in the steel disc were insulated with Rohacell bushes, with inner and outer diameters of 7.7 and 12.0 mm. Two total-temperature probes, see Fig. 4, were placed in diametrically opposite blade-cooling holes with the tip of each probe set back about 0.5 mm from the disc face. To reduce the blockage effect of the probes, the inner diameter of the Rohacell bushes was increased for these two blade-cooling holes. The probes were made from stainless-steel tubes, with inner and outer diameters of 2.8 and 3.2 mm, and two air vents of 0.5 mm diameter were drilled radially through each tube, as shown. The thermocouple wires were 0.10 mm diameter, and the 0.2 mm diameter bead was located downstream of the vents. Behind the bead, the tube was sealed with epoxy resin.

There are two significant problems in designing a probe to measure the total-temperature of the blade-cooling air. First, it is not easy to create solid-body rotation in swirling air flowing through a short, rotating hole, the axis of which is parallel to the axis of rotation. Meierhofer and Franklin [1] pointed out the problems they experienced while measuring the blade-coolant temperature in an adiabatic system. However, the probe-design described above was used successfully by El-Oun and Owen [2] to measure the total temperature of the blade-cooling air in an unheated pre-swirl rig, similar to the heated version used here. They estimated the uncertainty in their temperature measurements as ± 0.3°C, and their measured values were in good agreement with the theoretical temperatures for an adiabatic system. This suggests that the experimental uncertainties in measurements made with the probe used here should be small in an *adiabatic system*.

The second problem in temperature measurement occurs when heat transfer from the disc to the cooling air is significant: the Rohacell bushes in the blade-cooling holes reduce but do not cure this problem. As described in Section 4, computations show that there are large gradients of air temperatures in and around the blade-cooling holes, and it is unlikely that the probe described above will measure the true bulk-average temperature of the blade-cooling air. The uncertainties cannot be quantified.

3 Fluid Dynamics

3.1 Flow structure. As shown by Owen and Rogers [10], the flow structure in a rotating cavity depends principally on the inlet swirl ratio and the turbulent flow parameter, λ_T, where $\lambda_T = C_W \mathrm{Re}_\phi^{-0.8}$. It is helpful to note that $\lambda_T = \lambda_{T,fd} = 0.22$ corresponds to the nondimensional flow rate entrained by the boundary layer on one surface of a free-disc. From the authors' experience, relatively small flow rates ($\lambda_T < \lambda_{T,fd}$) are usually employed to cool the discs and seal the wheel spaces of gas-turbine engines; blade-cooling flow rates are usually much higher ($\lambda_T > \lambda_{T,fd}$).

Karabay et al. [7] showed that, at sufficiently large values of λ_T, free-vortex flow will occur in the core outside the boundary layers on the rotating surfaces. Under these conditions, $V_{\phi,\infty} \propto r^{-1}$, where $V_{\phi,\infty}$ is the tangential component of velocity in the core. As $V_{\phi,\infty} = \beta_p \Omega r_1$ at the outlet of the pre-swirl nozzles, where $r = r_1$, it follows that, for an ideal system with no losses,

$$\frac{V'_{\phi,\infty}}{\Omega r} = \beta_p x_1^2 x^{-2} \tag{1}$$

for $x \geqslant x_1$, where $x = r/b$ is the nondimensional radius and the dash is used to denote the ideal value.

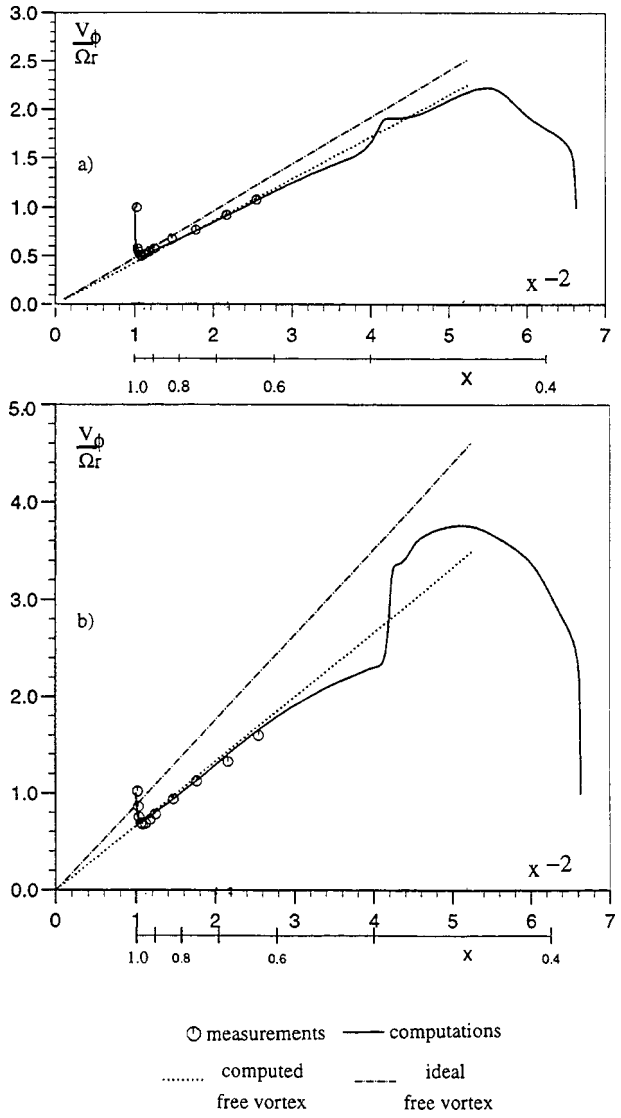


Fig. 5 Comparison between computed and measured variation of $V_\phi/\Omega r$ with x^{-2} for $\lambda_T=0.22$ and $Re_\phi=0.55\times10^6$: (*a*) $\beta_p=2.511$; (*b*) $\beta_p=4.535$.

In the cover-plate system considered here, where $\beta_p>1$, there is a nondimensional radius (x^*, say) at which $V'_{\phi,\infty}=\Omega r^*$. Hence, from Eq. (1),

$$x^*=\beta_p^{1/2}x_1. \tag{2}$$

Karabay et al. [7] showed that, provided $x^*<1$ (that is, $\beta_p x_1^2<1$), free-vortex flow will take place throughout the cavity when

$$\lambda_T\geqslant 0.437[1-(\beta_p x_1^2)^{1.175}]^{1.656}. \tag{3}$$

For smaller values of λ_T, nonentraining Ekman-type layers form in the outer part of the rotating cavity and free-vortex flow is confined to the inner part.

In practical cover-plate systems, there are losses between the outlet of the pre-swirl nozzles and the inlet to the rotating cavity. If $\beta_p>1$, viscous effects cause a loss of angular momentum of the fluid, reducing the effective swirl ratio. It is convenient to introduce an effective pre-swirl ratio, $\beta_{p,\mathrm{eff}}$, where, in comparison with Eq. (1),

$$\frac{V_{\phi,\infty}}{\Omega r}=\beta_{p,\mathrm{eff}}\,x_1^2x^{-2} \tag{4}$$

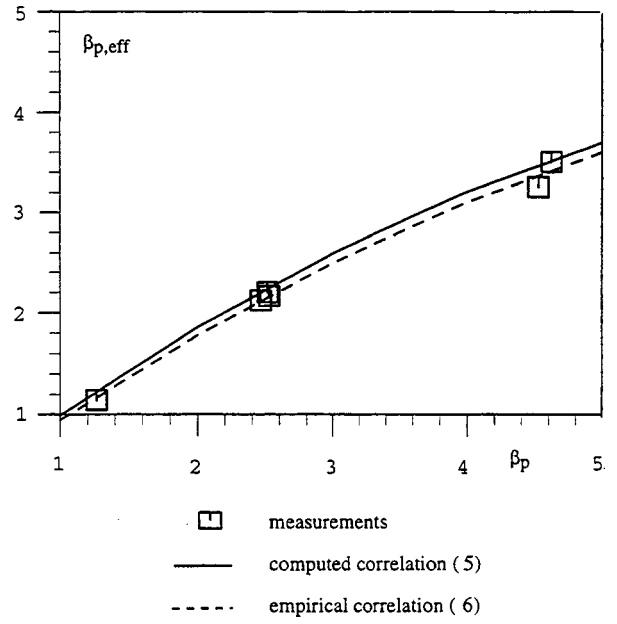


Fig. 6 Variation of $\beta_{p,\mathrm{eff}}$ with β_p for whole system

and, for $\beta_p>1$, $\beta_{p,\mathrm{eff}}<\beta_p$.

This relationship can be shown graphically by plotting $V_{\phi,\infty}/\Omega r$ versus x^{-2}: a free-vortex will appear as a straight line passing through the origin ($x^{-2}=0$); the gradient of this line can be used to calculate $\beta_{p,\mathrm{eff}}$. Figure 5 shows a comparison between the computed and measured variation of $V_{\phi,\infty}/\Omega r$ with x^{-2}. The results were obtained, for $\lambda_T=0.22$ and $Re_\phi=0.55\times10^6$, with $\beta_p=2.511$ and 4.535, using the axisymmetric computational model for the whole system and experimental apparatus described in Section 2. The agreement between the computations and measurements is very good, and there is clear evidence of free-vortex flow. Two straight lines are plotted: the higher one (the ''ideal free vortex'') corresponds to Eq. (1), and the lower one (the ''computed free vortex'') to Eq. (4). For the latter case, the line was arbitrarily forced through the computed value of $V_{\phi,\infty}/\Omega r$ at $x=0.67$ ($x^{-2}=2.22$). It can be seen that the difference between the two lines increases as β_p increases.

Karabay [12] found that the effects of Re_ϕ and λ_T on $\beta_{p,\mathrm{eff}}$ were negligible, and he proposed the following correlation for the computed values of $\beta_{p,\mathrm{eff}}$

$$\frac{\beta_{p,\mathrm{eff}}}{\beta_p}=1.053-0.062\beta_p. \tag{5}$$

A similar correlation was obtained for the experimental measurements where

$$\frac{\beta_{p,\mathrm{eff}}}{\beta_p}=1-0.056\beta_p. \tag{6}$$

The velocity measurements and computations were made over the range: $0.17<\lambda_T<0.35$, $0.5\times10^6<Re_\phi<1.5\times10^6$ and $1.1<\beta_p<4.6$. Figure 6 shows the good agreement between the two correlations and the experimental data, although it should be pointed out that the results were obtained for only one geometry (in which $r_1/b=0.43$).

The fact that $\beta_{p,\mathrm{eff}}<\beta_p$ when $\beta_p>1$ is attributed to the recirculation region, near the inlet to the rotating cavity, in which angular momentum is lost with a consequent reduction in swirl. The fact that the axisymmetric computations agree well with the measurements suggests that the three-dimensional effect of the pre-swirl nozzles is not significant. The losses are strongly related to β_p and only weakly to λ_T or Re_ϕ; the losses may, however,

depend on the system geometry, particularly if the cooling air passes through holes in the cover-plate rather than through an annular slot as considered here.

3.2 Pressure Distribution. Karabay et al. [11] showed that the radial distribution of pressure in the cavity can be estimated from the inviscid one-dimensional radial-momentum equation

$$\frac{1}{\rho}\frac{dp}{dr}=\frac{V_{\phi,\infty}^2}{r}-V_{r,\infty}\frac{dV_{r,\infty}}{dr}, \tag{7}$$

where $V_{\phi,\infty}$ can be approximated by Eq. (4) and $V_{r,\infty}$ from the continuity equation, where

$$V_{r,\infty}=\frac{\dot{m}}{2\pi\rho rs} \tag{8}$$

The local pressure coefficient, C_p, can be defined for compressible flow as

$$C_p=\frac{\gamma}{\gamma-1}\frac{p_1}{\frac{1}{2}\rho_1\Omega^2 r_1^2}\left\{\left(\frac{p}{p_1}\right)^{(\gamma-1)/\gamma}-1\right\}, \tag{9}$$

where, in the limit $p\to p_1$, the incompressible form is

$$C_p=\frac{p-p_1}{\frac{1}{2}\rho_1\Omega^2 r_1^2}. \tag{10}$$

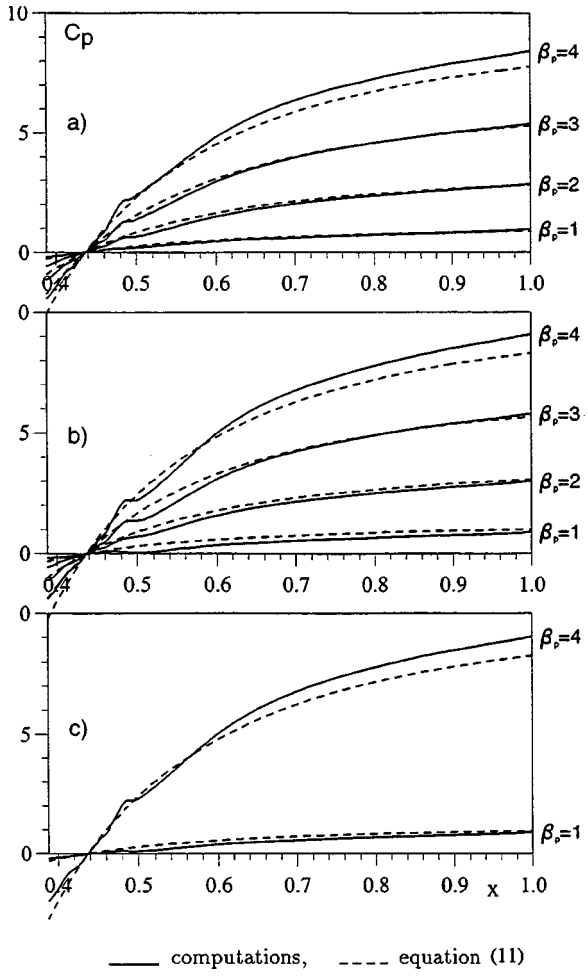


Fig. 7 Radial distribution of computed and theoretical pressure coefficients for whole system: (*a*) $Re_\phi=1.3\times10^6$, $\lambda_T=0.23$; (*b*) $Re_\phi=0.56\times10^6$, $\lambda_T=0.35$; and (*c*) $Re_\phi=1.3\times10^6$, $\lambda_T=0.35$.

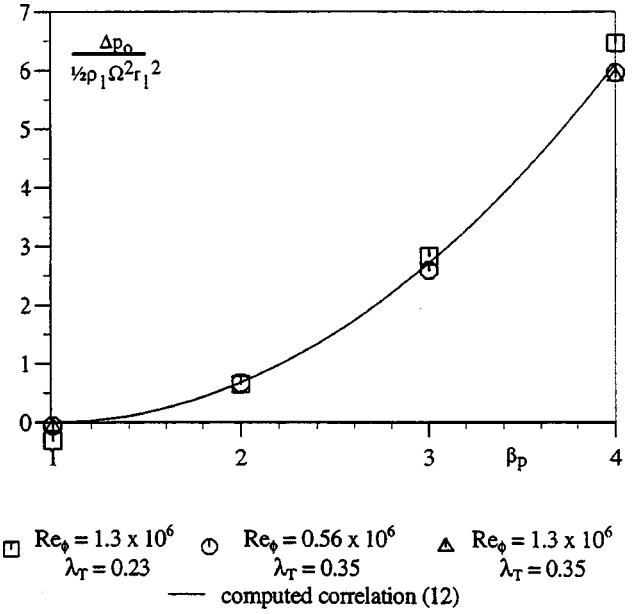


Fig. 8 Computed variation of nondimensional pressure loss β_p for whole system

Equation (7) can be integrated radially from $r=r_1$ to give

$$C_p=\left\{\beta_{p,\mathrm{eff}}^2+\left(\frac{\lambda_T}{2\pi G\,\mathrm{Re}_\phi^{0.2}}\right)^2\left(\frac{b}{r_1}\right)^4\right\}\left\{1-\left(\frac{r_1}{r}\right)^2\right\}. \tag{11}$$

Figure 7 shows comparisons between the computed values of C_p and Eq. (11) for the rig geometry (where $p=p_1$ at $r_1/b=0.43$ and $z/s=0.5$); Eq. (5) was used to calculate $\beta_{p,\mathrm{eff}}$. For $\beta_p\leqslant 3$, the agreement is good; for $\beta_p=4$, Eq. (11) slightly underestimates the computed pressure coefficients. Unfortunately, for the reason given in Section 2.2, no experimental data are available for comparison.

In an engine, the maximum pre-swirl ratio is limited by the available pressure upstream of the pre-swirl nozzles and the necessary pressure at inlet to the blade-cooling passages. The pressure losses in the pre-swirl nozzles and the blade-cooling passages are beyond the scope of this paper. However, the total pressures downstream of the pre-swirl nozzles (at $x=0.43$) and upstream of the blade-cooling passages (at $x=0.97$) were computed at $z/s=0.5$, in a stationary frame of reference, and the difference was used to define the loss in total pressure, Δp_0. This computed loss was correlated for $1<\beta_p<4$ by

$$\frac{\Delta p_0}{\frac{1}{2}\rho_1\Omega^2 r_1^2}=0.678(\beta_p-1)^2. \tag{12}$$

Figure 8 shows a comparison between this correlation and the computations; again there are no experimental data available. In practice, the loss would depend on the system geometry.

4 Thermodynamics and Heat Transfer

4.1 Thermodynamics of Pre-Swirl Process. Karabay et al. [7] used the steady-flow energy equation for an adiabatic pre-swirl system to show that

$$T_{t,2}-T_{0,1}=\frac{\Omega^2 r_2^2}{2c_p}\left(1-2\frac{r_1^2}{r_2^2}\beta_p\right), \tag{13}$$

where $T_{0,1}$ is the total temperature (in a *stationary frame* of reference) of the pre-swirl air (at $r=r_1$) and $T_{t,2}$ is the total temperature (in a *rotating frame*) of the blade-cooling air (at $r=r_2$); it is $T_{t,2}$ that controls the heat transfer inside the blade-cooling pas-

sages. It can be seen from Eq. (13) that $T_{t,2}$ decreases as β_p increases, with beneficial effects for blade cooling.

Karabay et al. [11] extended this adiabatic thermodynamic analysis to the case of a pre-swirl system where heat is transferred from the hot turbine disc to the cooling air; they ignored heat transfer from the cover-plate. Their result can be expressed as

$$T_{t,2}-T_{0,1}=\frac{\Omega^2 r_2^2}{2c_p}\left(1-2\frac{r_1^2}{r_2^2}\beta_p\right)+\pi\left(1-\frac{a^2}{b^2}\right)\times\frac{\mathrm{Nu}_{av}}{\mathrm{Pr}\,C_W}(T_s-T_{s,ad})_{av}. \quad (14)$$

Nu_{av} is the average Nusselt number defined as

$$\mathrm{Nu}_{av}=\frac{q_{s,av}b}{k(T_s-T_{s,ad})_{av}}, \quad (15)$$

and the subscript ''av'' refers to the radially weighted average value; $T_{s,ad}$ is the adiabatic disc-temperature, which is discussed below. Increasing β_p has a beneficial effect in the adiabatic case but what happens when heat transfer is significant? To answer this question, it is useful to use the Reynolds analogy.

4.2 The Reynolds Analogy Applied to Pre-Swirl Systems. Owen and Rogers [9] showed that the Reynolds analogy between angular momentum and heat transfer can be applied to boundary-layer flow in a rotating cavity under certain conditions. The most important of these conditions are that the Prandtl number of the fluid is unity (Pr=1), that the disc-temperature distribution is quadratic ($T_s \propto r^2$) and that the axial distributions of swirl and temperature are similar at entry to the cavity.

For the case of a rotating cavity with free-vortex flow throughout the core (see Section 3.1), Karabay et al. [11] showed that, for the simple rotating cavity, the adiabatic disc-temperature is given by

$$T_{s,ad}=T_{0,1}+\frac{\Omega^2 b^2}{2c_p}[Rx^2-\beta_p x_1^2(1+R)], \quad (16)$$

where R is the recovery factor, which for air can be approximated by $R=\mathrm{Pr}^{1/3}$. They also showed that, when the Reynolds analogy is valid,

$$\mathrm{Nu}_{av}=\frac{\mathrm{Re}_\phi C_m}{\pi\left(1-\frac{a^2}{b^2}\right)\left\{1+\frac{a^2}{b^2}-2\beta_p\frac{r_1^2}{b^2}\right\}}, \quad (17)$$

where C_m is the moment coefficient.

If Nu_{av} is to remain finite then C_m must equal zero when the denominator of Eq. (17) equals zero. This leads to the introduction of the *critical pre-swirl ratio*, $\beta_{p,\mathrm{crit}}$, such that $C_m=0$ when $\beta_p=\beta_{p,\mathrm{crit}}$, where

$$\beta_{p,\mathrm{crit}}=\frac{a^2+b^2}{2r_1^2}. \quad (18)$$

Although this result was derived using the Reynolds analogy, C_m does not depend directly on the Prandtl number of the fluid or on the temperature on the disc. Paradoxically, this implies that Eq. (18) is generally valid for all rotating cavities in which there is free-vortex flow throughout the core regardless of whether or not the Reynolds-analogy conditions are satisfied! Surprisingly, $\beta_{p,\mathrm{crit}}$ depends only on the geometric parameters (a, b, and r_1) and not on the flow parameters (Re_ϕ and λ_T).

Figure 9 shows the computed variation of C_m with β_p for two values of Re_ϕ and λ_T. The computations were carried out using the axisymmetric solver applied to the full pre-swirl system described in Section 2. For C_m, the moment was obtained by integrating the computed tangential component of shear stress from $r=80$ mm to $r=198.9$ mm (which were used for a and b in Eq. (18)); the pre-swirl radius, r_1 was 90 mm. These values give

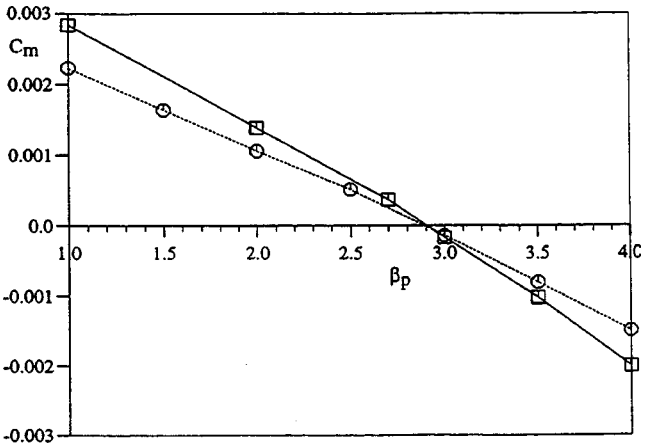


Fig. 9 Computed variation of moment coefficient with β_p for whole system (□ $\mathrm{Re}_\phi=0.56\times10^6$, $\lambda_T=0.35$; ○ $\mathrm{Re}_\phi=1.33\times10^6$, $\lambda_T=0.23$).

$\beta_{p,\mathrm{crit}}=2.837$ according to Eq. (18), and the computed values in Fig. 9 show that $C_m=0$ when $\beta_p=2.9$. For the simple rotating cavity considered by Karabay et al., the difference between the computed and theoretical values of $\beta_{p,\mathrm{crit}}$ were significantly smaller than the 2.2 percent difference found here for the whole system. In view of the complexity of the flow, the agreement between the computed and theoretical values of $\beta_{p,\mathrm{crit}}$ is remarkably good. Although the computations were conducted for only one geometry, there is no reason to believe that Eq. (18) will not be valid for other values of a, b, and r_1.

4.3 Average Nusselt Numbers. Karabay et al. [11] showed computationally that, for the simple cavity, $C_m=C_{m,\mathrm{fd}}$ when $\beta_p=0$, where $C_{m,\mathrm{fd}}$ is the free-disc moment coefficient, the value of which depends on Re_ϕ. They also showed that, when the Reynolds-analogy conditions were satisfied, there is *an optimal pre-swirl ratio*, $\beta_{p,\mathrm{opt}}$, where Nu_{av} is a minimum. The computed ratio $\beta_{p,\mathrm{opt}}/\beta_{p,\mathrm{crit}}$ was found to be less than unity for the Reynolds-analogy case; computations carried out at other conditions showed that the ratio depended on the flow parameters and on the distribution of temperature on the heated disc.

Figure 10 shows the variation of the computed values of Nu_{av} with β_p for the whole pre-swirl system described in Section 2. For $\mathrm{Re}_\phi=1.33\times10^6$ and $\lambda_T=0.23$, there is a minimum value of Nu_{av} at $\beta_p\approx2.5$, but the variation of Nu_{av} with β_p is very small. For $\mathrm{Re}_\phi=1.37\times10^6$ and $\lambda_T=0.35$, the minimum again occurs at $\beta_p\approx2.5$, but for this larger value of λ_T the variation of Nu_{av} is more pronounced. This is consistent with the computations of Karabay et al. for the simple cavity where the variation of Nu_{av} with β_p was found to increase as Re_ϕ and λ_T increased.

The magnitude of Nu_{av} has a direct effect on the temperature of the blade-cooling air, as Eq. (14) shows. The first term on the RHS of Eq. (14) is the adiabatic work term, which decreases as β_p increases; the second is the heat transfer term, which depends on Nu_{av}. For $\beta_p<\beta_{p,\mathrm{opt}}$, the work and heat transfer terms are aligned: $T_{t,2}$ decreases as β_p increases. For $\beta_p>\beta_{p,\mathrm{opt}}$, the two terms are opposed: whether $T_{t,2}$ increases or decreases as β_p increases depends on the relative magnitude of the terms.

The above computations were made with the axisymmetric solver. In the experimental rig, and in a gas turbine, the flow and heat transfer are three-dimensional in the vicinity of the blade-cooling holes. This has a significant effect on the temperature of the blade-cooling air, as discussed below.

4.4 Temperature Rise of Blade-Cooling Air. Karabay [12] used the axisymmetric solver described in Section 2 to compute the bulk-average value of ΔT (where $\Delta T=T_{t,2}-T_{0,1}$) for a range of values of Re_ϕ, λ_T, and β_p. The computed values of ΔT

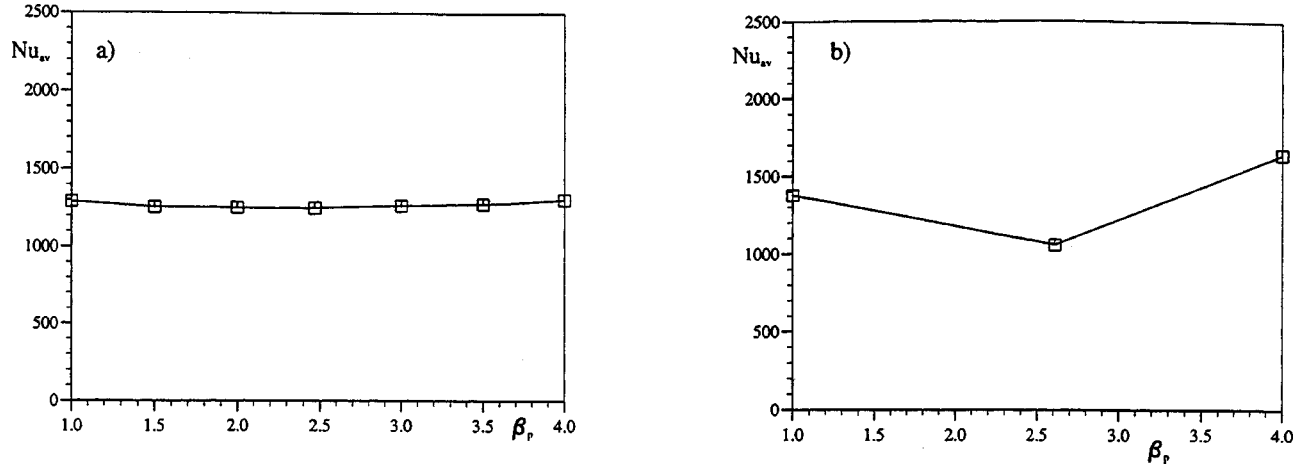


Fig. 10 Computed variation of Nu_{av} with β_p for whole system: (*a*) $Re_\phi=1.33\times10^6$, $\lambda_T=0.23$; (*b*) $Re_\phi=1.37\times10^6$, $\lambda_T=0.35$.

consistently underestimated the measurements: on average, the measured values were 50 percent higher than the computations.

Karabay [12] also describes three-dimensional computations (carried out by Dr. J. X. Chen) of the flow and heat transfer in the simple cavity. These computations were made with a turbulent version (incorporating the Launder-Sharma [13] turbulence model) of the three-dimensional code used by Chen et al. [3,4]. They were conducted for a six-degree segment of the rotating cavity, corresponding to the region around one of the 60 blade-cooling holes in the rotor of the experimental rig. A $65\times60\times11$ (axial×radial×tangential) grid was used, with 72 grid points in the blade-cooling hole. The thermal boundary conditions were similar to those used for the axisymmetric solver, and cyclic symmetry in the tangential direction was assumed on the two radial boundaries.

Figure 11 shows comparisons between computed and measured local Nusselt numbers for five test cases. The two-dimensional computations for the whole system were made using the axisymmetric solver described in Section 2; additional measurements and two-dimensional computations of Nu for the whole system are presented by Pilbrow et al. [8]. The two-dimensional and three-dimensional computations made by Chen were only for the simple cavity, so as to reduce the computational complexity; for the two-dimensional cases, the blade-cooling holes were modeled as an annular equivalent-area slot. Pilbrow et al. [8] produced better agreement between their axisymmetric computations, using the Morse turbulence model, and the measured Nusselt numbers than that shown in Fig. 11, where the Launder-Sharma model was used. Despite the superiority of the Morse model, the Launder-Sharma version was used here for comparison with Chen's computations.

Most of the computed Nusselt numbers shown in Fig. 11 underestimate the measured values. The two-dimensional and three-dimensional computations of Chen, for the simple cavity, are in close agreement over most of the cavity: the largest differences occur in the vicinity of the blade-cooling holes at $x=0.97$. The computations for the simple cavity differ from those for the whole system, particularly at the smaller radii where the computed flows are significantly different.

In the r-ϕ plane, the three-dimensional computations produced very large tangential and radial gradients of both Nusselt numbers and air temperatures near the blade-cooling holes, with the highest temperatures close to the perimeter of the hole and the lowest near the center. Table 1 shows a comparison between the computed and measured values of ΔT for six cases, which were also studied experimentally: $\Delta T(a)$ corresponds to the two-dimensional computations for the whole system, and $\Delta T(b)$ and $\Delta T(c)$ correspond respectively to the two-dimensional and three-dimensional com-

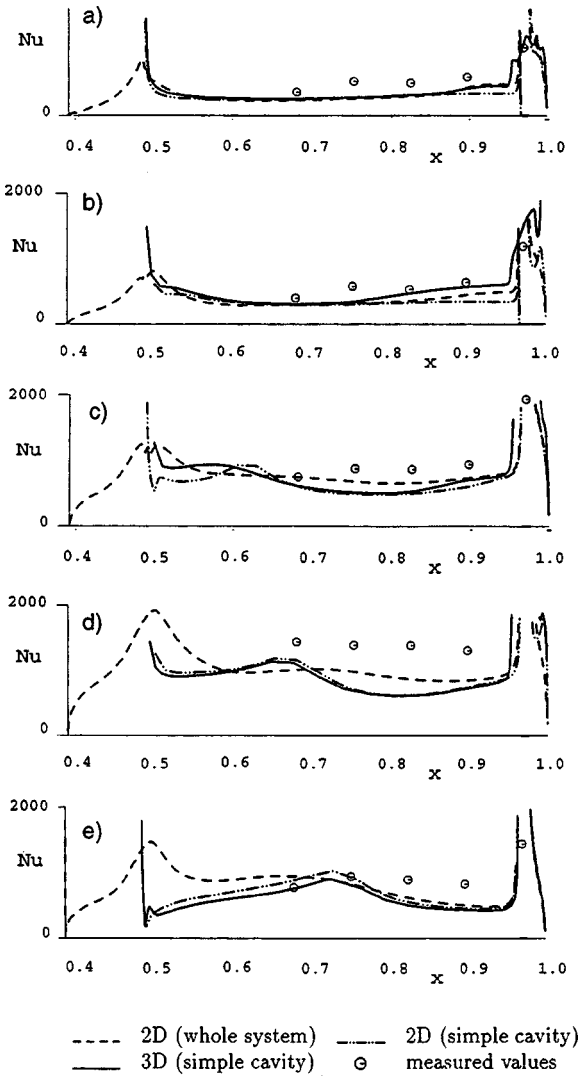


Fig. 11 Comparison between computed and measured radial variation of Nu with x: (*a*) $Re_\phi=0.535\times10^6$, $\lambda_T=0.173$, $\beta_p=1.110$, (*b*) $Re_\phi=0.542\times10^6$, $\lambda_T=0.176$, $\beta_p=1.537$, (*c*) $Re_\phi=0.898\times10^6$, $\lambda_T=0.351$, $\beta_p=2.049$, (*d*) $Re_\phi=0.965\times10^6$, $\lambda_T=0.349$, $\beta_p=2.866$, (*e*) $Re_\phi=0.588\times10^6$, $\lambda_T=0.353$, $\beta_p=3.059$.

Table 1 Comparison between computed and measured values of ΔT(°C): (*a*) two-dimensional, whole system; (*b*) two-dimensional, simple cavity; and (*c*) three-dimensional, simple cavity.

β_p	$Re_\phi \times 10^6$	λ_T	*(a)* ΔT	*(b)* ΔT	*(c)* ΔT	Measured ΔT
1.110	0.535	0.173	12.2	13.2	18.5	21.9
1.267	1.490	0.175	11.9	10.6	15.6	16.2
1.537	0.542	0.176	11.4	11.8	15.1	18.0
2.049	0.898	0.351	6.6	7.3	10.3	10.6
2.866	0.965	0.349	6.2	6.0	7.5	6.5
3.059	0.588	0.353	6.6	7.2	7.9	7.1

putations of Chen for the simple cavity. For most tests, the two-dimensional computations ($\Delta T(a)$ and $\Delta T(b)$) significantly underestimate the measured values, as was noted above, whereas the three-dimensional computations ($\Delta T(c)$) are closer to the measured values. It should be pointed out that, as a consequence of the large velocity and temperature gradients near the blade-cooling holes, the isotropic assumption of the gradient-diffusion modeling of the turbulence is likely to be inaccurate.

In view of the uncertainties in both the computations and the measurements, these results cannot be regarded as conclusive. However, they do strongly suggest that three-dimensional effects have a significant influence on both the heat transfer around the blade-cooling holes and the temperature rise of the cooling air. As discussed in Section 2.2.2, these results underline the need for experimentalists to take great care in how they measure the temperature of the blade-cooling air.

5 Conclusions

A combined theoretical, computational, and experimental study has been undertaken of the flow and heat transfer in a cover-plate pre-swirl system. The results support and extend those presented in earlier studies, and the following conclusions are relevant to the designer of internal cooling-air systems.

- For sufficiently large flow rates, free-vortex flow occurs in the core of fluid outside the boundary layers on the rotating surfaces of the cavity between the cover-plate and the rotor.
- The free-vortex flow is specified by an effective pre-swirl ratio, $\beta_{p,\mathrm{eff}}$, which is smaller than the inlet pre-swirl ratio, β_p, immediately downstream of the pre-swirl nozzles. There is a simple correlation for $\beta_{p,\mathrm{eff}}$, supported by experimental evidence for one particular geometry, and the ratio of $\beta_{p,\mathrm{eff}}/\beta_p$ decreases as β_p increases.
- A theoretical expression for the pressure coefficient, C_p, is in good agreement with computed values for $\beta_p \leqslant 3$. Also, a correlation, based on computed pressures in the system, has been produced relating the loss of total pressure to the pre-swirl ratio. No pressure measurements are available to validate these results.
- It has been shown, theoretically and computationally, that the moment coefficient for the rotor, C_m, decreases as β_p increases. There is a critical pre-swirl ratio, $\beta_{p,\mathrm{crit}}$, for which $C_m = 0$; for $\beta_p > \beta_{p,\mathrm{crit}}$, $C_m < 0$.
- It has been shown computationally that there is an optimal pre-swirl ratio, $\beta_{p,\mathrm{opt}}$, for which the average Nusselt number, Nu_{av}, is a minimum.
- It has been shown theoretically and computationally, using the energy equation, that the total temperature of the blade-cooling air, T_{t2}, depends strongly on β_p: T_{t2} decreases as β_p increases providing $\beta_p < \beta_{p,\mathrm{opt}}$; for $\beta_p > \beta_{p,\mathrm{opt}}$, whether T_{t2} increases or decreases as β_p increases depends on the relative magnitudes of the heat transfer and work terms in the energy equation.
- Computed values of T_{t2} have been compared with experimental measurements. On average, measured increases of air temperature are around 50 percent higher than values obtained with axisymmetric solvers. Three-dimentional computations, giving results much closer to the measured values, show that there are large gradients of air temperature inside the blade-cooling holes. The three-dimentional computations serve to emphasize the need for care in making total-temperature measurements inside blade cooling holes.

In summary, the theoretical expressions for $V_{\phi,\infty}/\Omega r$, C_p, $(T_{t,2} - T_{0,1})$, $(T_{s,ad} - T_{0,1})$, and $\beta_{p,\mathrm{crit}}$ given in Eqs. (4), (11), (14), (16), (18), respectively are expected to apply to other geometries and to other flow parameters as well as to those tested here. However, some quantities, such as the pressure losses and Nusselt numbers, depend strongly on the geometry and on the flow parameters of the individual system. Although this paper has presented a number of concepts and has given theoretical models for some important design parameters, more computational and experimental work is needed to quantify these parameters.

Acknowledgments

The authors thank the UK Engineering and Physical Sciences Research Council and ALSTOM Gas Turbines Ltd. (formerly European Gas Turbines Ltd.) for funding this research project. The three-dimensional computations were conducted by Dr. J. X. Chen while he was at the University of Bath.

Nomenclature

a = inner radius of cavity
b = outer radius of cavity
C_m = moment coefficient ($= M/\frac{1}{2}\rho\Omega^2 b^5$)
C_p = pressure coefficient ($= (p - p_1)/\frac{1}{2}\rho\Omega^2 r_1^2$)
c_p, c_v = specific heat at constant pressure and constant volume, respectively
C_w = nondimensional mass flow rate ($= \dot{m}/\mu \mathrm{b}$)
G = gap ratio ($= s/b$)
k = turbulent kinetic energy; thermal conductivity
M = moment for one side of the rotor
$\dot{m}$ = mass flow rate of blade-cooling air
Nu = local Nusselt number $[= rq_s/k(T_s - T_{s,ad})]$
Nu_{av} = average Nusselt number $[= bq_{s,av}/k(T_s - T_{s,ad})_{av}]$
p = static pressure
Pr = Prandtl number ($= \mu c_p/k$)
Pr_t = turbulent Prandtl number
q = heat flux
q_s = convective heat flux from disc to air
r, ϕ, z = radial, tangential and axial coordinates
R = recovery factor
Re_ϕ = rotational Reynolds number ($= \rho\Omega b^2/\mu$)
s = axial width of rotating cavity
S = axial width of whole system
T = temperature
U_τ = friction velocity ($= \sqrt{\tau_w/\rho}$)
V_r, V_ϕ, V_z = time-averaged radial, circumferential, axial components of velocity in a stationary frame
V'_ϕ = ideal value of V_ϕ in a free vortex
x = nondimensional radial coordinate ($= r/b$)
y = distance normal to the wall
y^+ = nondimensional distance ($= \rho y U_\tau/\mu$)
z = axial distance
β_p = pre-swirl ratio ($= V_\phi/\Omega r$ at $r = r_1$)
Δp_0 = loss in total pressure from pre-swirl nozzles to blade-cooling passages
ΔT = total-temperature difference ($= T_{t,2} - T_{0,1}$)
ε = turbulent energy dissipation rate
γ = ratio of specific heats ($= c_p/c_v$)
λ_T = turbulent flow parameter ($= C_w/\mathrm{Re}_\phi^{0.8}$)
μ, μ_t = dynamic and turbulent viscosities
ρ = density
Ω = angular speed of rotor

Subscripts

- ad = adiabatic value
- av = radially-weighted average
- b = blade-cooling air
- crit = value of β_p when $C_m = 0$
- eff = effective value (of β_p)
- fd = free-disc value
- o = total value in stationary frame of reference
- opt = value of β_p when Nu_{av} is minimum
- p = pre-swirl air
- s = disc surface
- t = total value in rotating frame of reference
- ∞ = value in core outside boundary layers
- 1 = inlet to system (at radial location of pre-swirl nozzles)
- 2 = outlet from system (at radial location of blade-cooling holes)

Superscripts

- * = value at stagnation point (where $V_{\phi,\infty} = \Omega r$)

References

[1] Meierhofer, B., and Franklin, C. J., 1981, ''An Investigation of a Pre-swirled Cooling Airflow to a Gas Turbine Disk by Measuring the Air Temperature in the Rotating Channels,'' ASME Paper 81-GT-132.

[2] El-Oun, Z., and Owen, J. M., 1989, ''Pre-Swirl Blade-Cooling Effectiveness in an Adiabatic Rotor-Stator System,'' ASME J. Turbomach., **111**, pp. 522–529.

[3] Chen, J., Owen, J. M., and Wilson, M., 1993, ''Parallel-Computing Techniques Applied to Rotor-Stator Systems: Fluid Dynamics Computations'' in *Numerical Methods in Laminar and Turbulent Flow*, **8**, Pineridge Press, Swansea, pp. 899–911.

[4] Chen, J., Owen, J. M., and Wilson, M., 1993, Parallel-Computing Techniques Applied to Rotor-Stator Systems: Thermal Computations, in *Numerical Methods in Thermal Problems*, **8**, Pineridge Press, Swansea, pp. 1212–1226.

[5] Popp, O., Zimmermann, H., and Kutz, J., 1996, ''CFD Analysis of Cover-Plate Receiver Flow,'' ASME Paper 96-GT-357.

[6] Wilson, M., Pilbrow, R., and Owen, J. M., 1997, ''Flow and Heat Transfer in a Pre-Swirl Rotor-Stator System,'' ASME J. Turbomach., **119**, pp. 364–373.

[7] Karabay, H., Chen, J. X., Pilbrow, R., Wilson, M., and Owen, J. M., 1999, ''Flow in a Cover-Plate Pre-Swirl Rotor-Stator System,'' ASME J. Turbomach., **121**, pp. 160–166.

[8] Pilbrow, R., Karabay, H., Wilson, M., and Owen, J. M., 1999, ''Heat Transfer in a 'Cover-Plate' Pre-Swirl Rotating-Disc System,'' ASME J. Turbomach., **121**, pp. 249–256.

[9] Owen, J. M., and Rogers, R. H., 1989, *Flow and Heat Transfer in Rotating Disc Systems: Vol. 1, Rotor-Stator Systems*, Research Studies Press, Taunton, UK (John Wiley, New York).

[10] Owen, J. M., and Rogers, R. H., 1995, *Flow and Heat Transfer in Rotating Disc Systems: Vol. 2, Rotating Cavities*, Research Studies Press, Taunton, UK (John Wiley, New York).

[11] Karabay, H., Wilson, M., and Owen, J. M., 1999, ''Predicting Effects of Swirl on Flow and Heat Transfer in a Rotating Cavity,'' submitted to Int. J. Heat Fluid Flow.

[12] Karabay, H., 1998, ''Flow and Heat Transfer in a Cover-Plate Pre-Swirl Rotating-Disc System,'' Ph.D. thesis, University of Bath, UK.

[13] Launder, B. E., and Sharma, B. I., 1974, ''Application of the Energy Dissipation Model of Turbulence to the Calculation of the Flow Near a Spinning Disc,'' Letters in *Heat and Mass Transfer*, pp. 131–138.

[14] Morse, A. P., 1988, ''Numerical Prediction of Turbulent Flow in Rotating Cavities,'' ASME J. Turbomach., **110**, pp. 202–215.

[15] Chen, J., Gan, X., and Owen, J. M., 1996, ''Heat Transfer in An Air-Cooled Rotor-Stator System,'' ASME J. Turbomach., **118**, pp. 444–451.

Stress Relaxation Testing of Service Exposed IN738 for Creep Strength Evaluation

David A. Woodford
Materials Performance analysis, Inc.,
1707 Garden Street,
Santa Barbara, CA 93101

Standard size and miniature specimens of IN738 were taken from a service exposed turbine blade and vane for comparative stress relaxation testing at 800C, 850C, and 900C. Base data taken from root section material were used to construct stress versus creep rate parametric curves which could be used directly in design. Up to five decades in creep rates were obtained at each temperature from tests lasting less than one day. The data were also presented in the form of stress versus predicted times to 0.5 percent creep which compared well with available long time creep data. Differences were noted in specimens taken from different locations in the airfoil regions which probably resulted from differences in grain size or orientation. Based on these measurements it was concluded that there was no significant effect of section size on creep strength as defined by this test, and that the alloy was quite insensitive to prior deformation and thermal exposures. A life management procedure, using a combination of creep strength evaluation based on the stress relaxation test and a separate fracture evaluation measurement, is outlined in which end of useful life is defined in terms of minimum acceptable performance levels. [S0742-4795(00)01803-2]

Introduction

Design for Performance is a recently developed methodology for evaluating the creep strength and fracture resistance of high temperature materials [1]. Whereas the traditional approach involves long time testing and attempts to incorporate microstructural evolution in the test measurements, the new approach aims to exclude these changes in a short time high precision test. However, the test may be used unambiguously to evaluate the consequences of such changes in service exposed samples. Also, the traditional approach uses a single test to evaluate both creep strength (e.g., minimum creep rate) and fracture resistance (the time to rupture). In practice, because there is generally a good inverse relationship between minimum creep rate and time to rupture in ductile materials, they are in fact separate measures of the same property, i.e., creep strength. *Design for performance* recognizes that different tests are necessary to measure the two properties. For creep strength, a stress versus creep rate response is determined from a stress relaxation test (SRT), and for fracture resistance a constant displacement rate test (CDR) of a notched or smooth tensile specimen is performed at a temperature where the part is most vulnerable to fracture.

Although the deformation histories are quite different, for many high temperature materials the two approaches may give closely similar results for creep strength under certain testing conditions. However, in some cases there are significant differences because very specific microstructural evolution occurs during long time testing at a fixed stress. However, unless the service application involves a similar deformation history, i.e., long times at constant stress and temperature, it may not be appropriate to use the long term data as a standard with which to compare all other testing.

The *design for performance* methodology has been used successfully for metals [2,3], polymers [4,5] and ceramics [6,7]. It has been applied to accelerate and optimize materials development, provide a basis for design analysis, and offer a framework for remaining life evaluation of operating components. However, it is important ultimately to judge its potential in terms of an objective assessment of its cost and effectiveness relative to high temperature component performance. This is very difficult because of the natural tendency to judge a new method in terms of the measurements made using the traditional method, and based on a quantitative comparison between the tests.

In this paper we will show a comparison of creep strength measurements for IN738 using the two methods. Similarities and differences will be noted, and possible approaches to use the SRT data directly in alloy optimization, design analysis and life assessment will be presented. Included in this study were effects of temperature and strain, and evaluation of repeatability and section size effects.

Experimental Procedure

One engine run blade and vane (48,000 engine operating hours) were used in this program. Standard tensile specimens were taken from the blade shank, 76 mm long with a gage section of 25.4 by 4.1 mm diameter. These were used for the SRT base testing to represent undamaged material. Specimens $5L$, $6L$, and $7L$ were tested at 800C, 850C, and 900C respectively. Each specimen was relaxed from 0.2 percent, 0.4 percent, 0.8 percent, and 1.5 percent total strain. These results provided a basis for constructing pseudo stress-strain curves as a function of creep rate, and also a comparison basis with long time creep data to 0.5 percent creep strain. This approach is consistent with previous studies in other alloys [2,3].

In addition, miniature specimens 41.3 mm long with a gage section of 25.4 by 2 mm diameter were taken from the same location as the larger specimens in the shank, and also from the blade and vane airfoil sections. For the blade, specimens $1b$, $2b$, and $3b$ were all close to the pressure face and located near the leading edge, middle and trailing edge, respectively. For the vane, specimens $2v$ and $3v$ were on the pressure face near the leading and trailing edges, $1v$ was from a thin bridge across the leading edge, and $4v$ was from the thickest section of the convex side.

The test procedure involved loading to the prescribed strain level, then holding the strain constant for twenty hours during which the stress relaxed as elastic strain was replaced with inelastic creep strain. Using the measured elastic modulus on loading,

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-285. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

the stress versus time response was converted to a stress versus creep rate curve covering approximately five decades in creep rate, as described previously [1]. This curve is the primary product of the analysis, and provides a comprehensive measure of the current creep strength of the material. For loading strains above about 0.4 percent (i.e., about 0.1 percent plastic strain), the plastic deformation is expected to lead to some recovery processes occurring during the test. To conform with traditional analysis and design approaches, the data were further reduced to produce pseudo-stress-strain curves as a function of creep rate, and stress versus pseudo-time to 0.5 percent creep strain. In the latter case, since the prediction is for several thousand hours, based on the initial strain and the very low creep rates obtained during relaxation in a twenty hour test, it is appropriate to use the term pseudo-time. Again, we emphasize that the SRT test does not contain information on time-dependent microstructural changes beyond the test duration. Nevertheless, a loading strain of 0.8 percent ($\sim$0.5 percent plastic strain) is used for comparison with traditional creep tests to 0.5 percent creep strain. Perfect agreement is unlikely because strain is not a state variable. However, this comparison serves as a legitimate calibration of the two methods.

After 20 h the normal procedure was to unload to a low stress ($\sim$5 MPa) and hold that stress constant for at least two hours. Most materials show a significant amount of anelastic creep recovery, i.e., contraction. This is because the creep strain consists of both non-recoverable time-dependent plastic strain and recoverable time-dependent anelastic strain.

Results

Standard Specimens. Figure 1 is an example stress-strain plot for the SRT runs at 850C in specimen 6*L*. The four successive runs are plotted in terms of stress versus log time in Fig. 2 and fitted to fourth order polynomials. The form of these curves allowed easy differentiation, and the creep rate could then be obtained by dividing the stress rate by the elastic modulus measured on loading. The values used for E were 145,000 (5*L*), 140,000 (6*L*), 140,000 (7*L*), and 135,000 MPa (7*L* rpt). These are averages of the values measured for each loading at 800, 850, 900C, and 900C, respectively. The values were typically repeatable to within 10 percent which correspondingly leads to a variation in creep rate and estimated creep time of 10 percent. A plot of log stress versus creep rate for the four runs at 850C is shown in Fig. 3. Pseudo stress strain curves generated from a cross plot of Fig. 3 are shown in Fig. 4. These may be used in design in a similar way to isochronous stress-strain curves cross plotted from traditional creep curves. Using the SRT data, however, the design limits would preferably be set in terms of creep rate rather than time.

A similar sequence of curves was generated for specimen 5*L* at 800C and 7*L* at 900C. Additionally, since the last run was omitted in 7*L*, a complete repeat series was conducted (7*L* rpt). This allowed useful direct comparison of duplicate tests. The shapes of the stress-creep rate curves were closely duplicated, and the stress levels are similar.

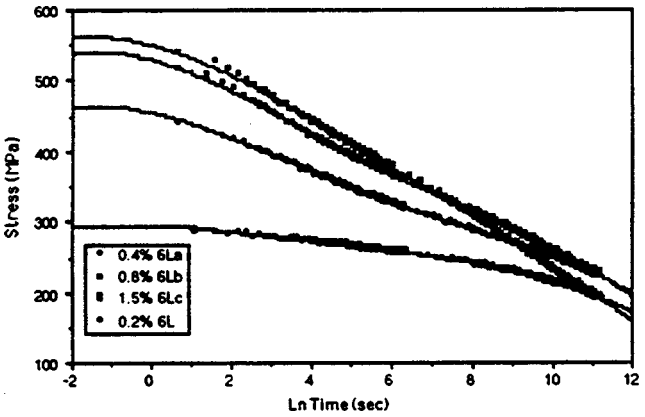


Fig. 2 Stress versus Ln time at 850C for specimen 6*L*

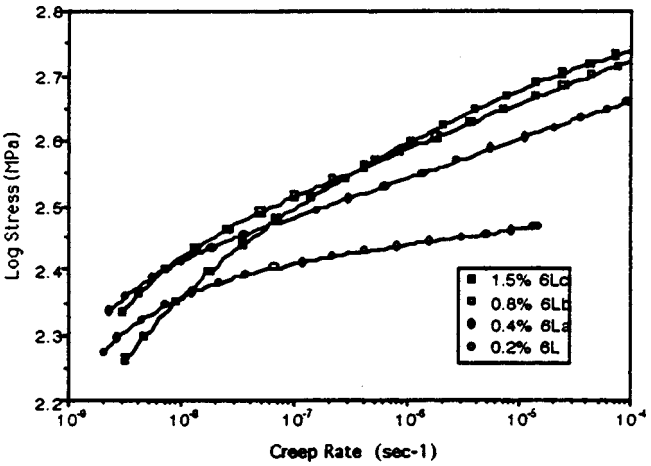


Fig. 3 Stress versus creep rate at 850C for specimen 6*L*

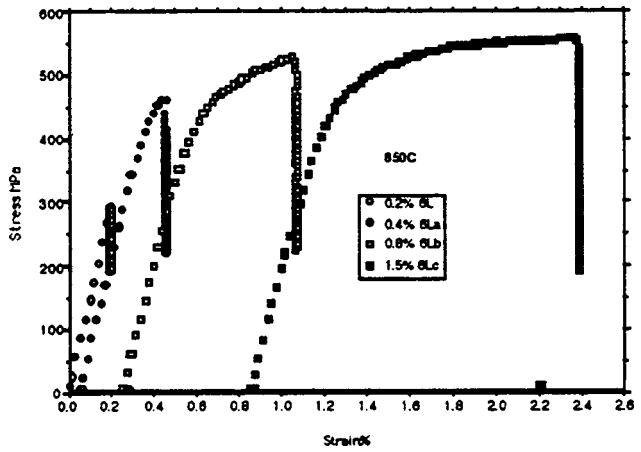


Fig. 1 Stress-strain curves for specimen 6*L* at 860C

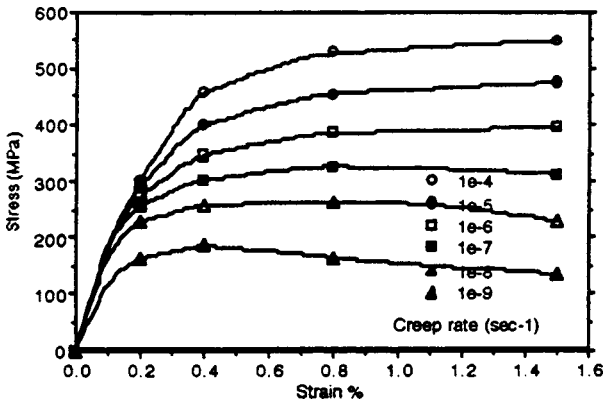


Fig. 4 Pseudo stress-strain curves at 850C as a function of creep rate

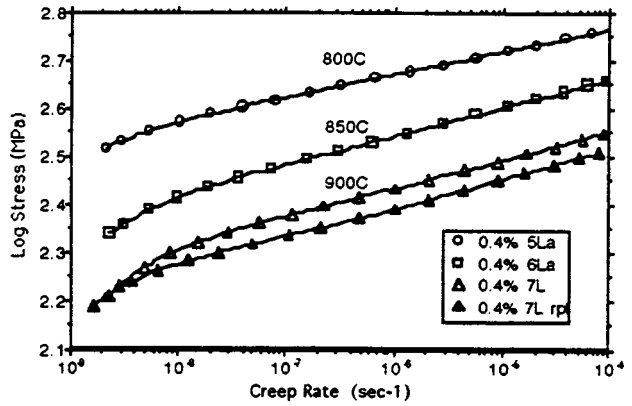


Fig. 5 Stress versus creep rate curves from 0.4 percent strain as a function of test temperature

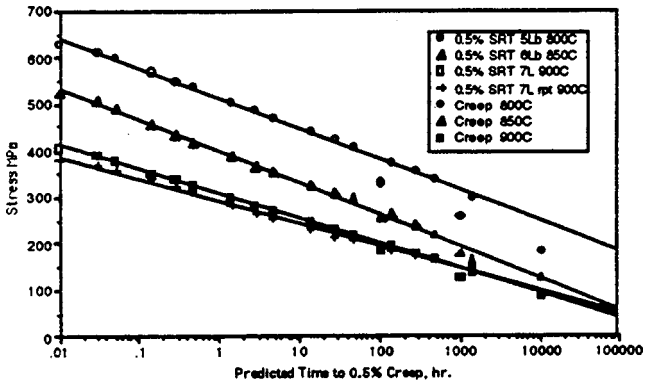


Fig. 7 Stress versus predicted time to 0.5 percent creep compared with actual long time creep data

For SRT runs with appreciable plastic strain prior to starting the test, there was a tendency to crossover at low strain rates. This results from dynamic recovery processes which also normally occur during a creep test. Although the deformation history is different, the SRT run from 0.8 percent total strain (~0.5 percent plastic strain) provides the most logical comparison with 0.5 percent conventional creep data. To minimize microstructural changes during the SRT run, but also characterize the current creep strength of the alloy, the run from 0.4 percent total strain (~0.1 percent plastic strain) is the appropriate choice. These two sequences were used in all comparisons. It is important to note that in all these tests there appears to be no significant memory after relaxation of previous deformation and thermal exposures for this alloy in this range of test temperatures; even after appreciable plastic deformation the yield stresses were the same and the flow curves closely similar during the test sequence (see the example at 850C in Fig. 1). Thus, differences among the specimens should be retained during a test sequence.

Figure 5 shows the stress-creep rate curves from 0.4 percent at the three temperatures. Using an intercept at log stress = 2.5, and an average creep rate for the duplicate tests at 900C, it was possible to develop an Arrhenius type scaling law from Fig. 5. This allowed a master plot to be constructed covering 10 decades in creep, rate and extrapolation of the lower temperature data. Thus, Fig. 6 shows possible design points at 800C and 850C based on creep rates of $3\times10^{11}\,s^{-1}$ (1 percent in 100,000 h).

Using the 0.8 percent SRT tests, a pseudo time was calculated for various creep rates and 0.5 percent creep, and plotted as predicted times in Fig. 7 for the three temperatures. For example, the lowest creep rate of $10^{-9}\,s^{-1}$ gives a pseudo time of 1390 h. The resulting curves are shown as linear plots on a semi-logarithmic scale in Fig. 7. Also shown in the figure are times to 0.5 percent creep for IN738 taken from MPa files. Since these were different heats the comparison has limited value. However, it is generally consistent with closer similarity at higher temperatures. This has been observed previously and will be discussed subsequently.

Miniature Specimens. All SRT tests were run at 850C from 0.4 percent and 0.8 percent total strain. The vane specimens, $1v-4v$, had average elastic moduli of 133,000, 138,000, 125,000, and 124,000 MPa, respectively. Consistent with the modulus differences, specimens $1v$ and $2v$ were stronger than $3v$ and $4v$ (see Fig. 8).

Figure 9 shows the stress-strain curves for blade specimens $1b-4b$ with calculated elastic moduli of 120,000, 130,000, 122,000, and 140,000 MPa, respectively. In this case $1b-3b$ from the airfoil have similar strengths, with $3b$ slightly higher. However, $4b$ taken from the blade shank, adjacent to the standard specimens, had a much higher modulus and strength. Figure 10 compares the predicted times to 0.5 percent creep for the three airfoil specimens and confirms the slightly higher strength for specimen $3b$. Also included are actual times to 0.5 percent creep taken from a separate source for miniature specimens from a similar engine run blade. Specimen IBSJ fell well below the SRT lines but was observed to have significant casting pores. Specimen 1BbJ was close to the SRT predictions.

A comparison of the blade and vane miniature specimens in Fig. 11 in terms of stress-creep rate from 0.4 percent strain at 850C shows that $1v$ and $2v$ have significantly higher creep strength than all the other airfoil specimens. However, miniature

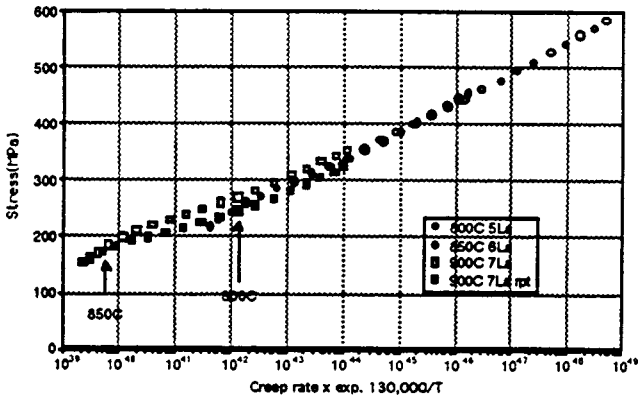


Fig. 6 Stress versus parameter plot showing possible design points at 800C and 850C

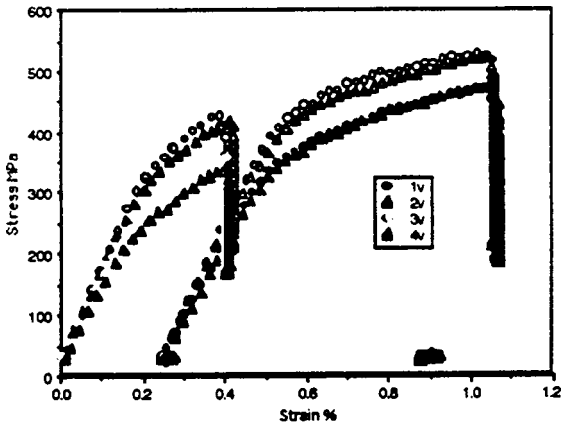


Fig. 8 Stress-strain curves at 850C for miniature specimens from vane

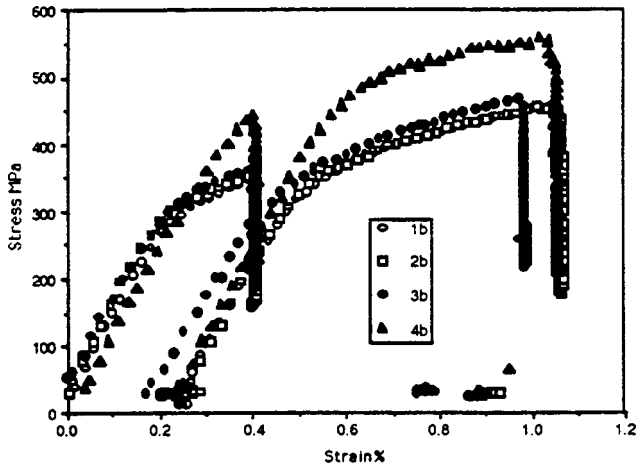


Fig. 9 Stress-strain curves at 850C for miniature specimens from blade

specimen $4b$ and standard specimen $6L$ taken from the blade shank are both a little stronger than $1v$ and $2v$. The two bands of data in Fig. 11 are separated by roughly an order of magnitude in creep rates at a given stress at 850C. It would be of great interest to characterize these differences in terms of the microstructures. This may be a response to a larger grain size and dendrite arm spacing in the shank area of the blade and the leading edge area of the blade. However, it is possible that it is primarily an orientation

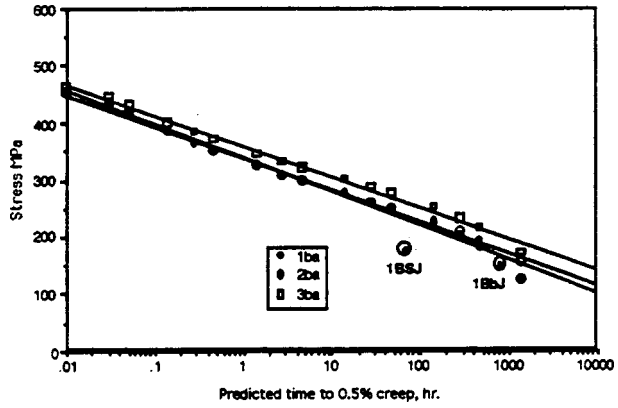


Fig. 10 Stress versus predicted time to 0.5 percent creep for blade specimens

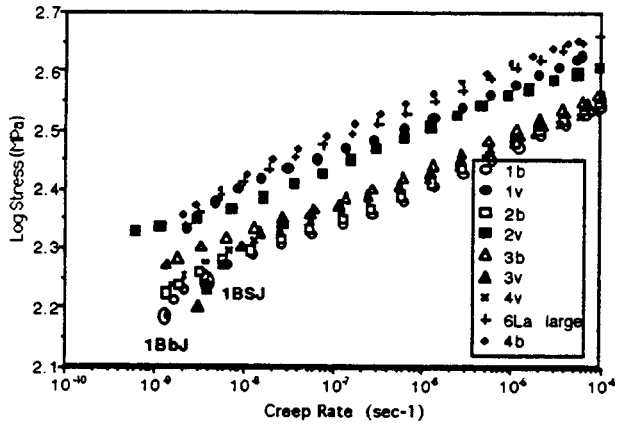


Fig. 11 Stress versus creep rate for all specimens at 850C

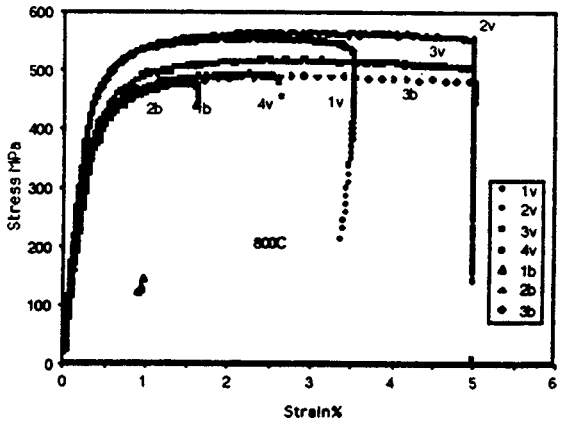


Fig. 12 Tensile tests to failure of miniature specimens

effect. The higher elastic modulus associated with the higher strength also suggests that local grain orientations make a major contribution to creep strength variation. The two minimum creep rate points fall on the lower bound of the data.

All specimens showed a significant creep recovery after unloading and holding at a small constant stress. This anelastic effect, apparent in Figs. 1, 8, and 9 amounted to between 10 and 20 percent of the inelastic strain accumulated during each SRT run.

The CDR tests were inadvertently run under machine control rather than extensometer strain control. The result is that the specimens did not unload to zero stress during failure. However, the elongation at failure is not appreciably affected by the control mode, and clear distinction among the specimens could be made. Figure 12 shows the data plotted as nominal stress versus nominal strain. Specimens $2v$, $3v$, and $3b$ reached the extensometer limit at 5 percent. At the other extreme, specimens $1b$, $2b$, $4b$, and $1v$ failed after about 1.5 percent, 1 percent, 1 percent, and 3.5 percent, respectively. However, $4b$ is not included in Fig. 12 since it failed outside the gage section so there is the possibility of a casting defect. Although these specimens had previously been deformed about 1 percent plastic strain during the SRT tests, they may be considered to have poor fracture resistance. The fracture results show no connection with the creep strength comparisons for all the specimens.

Discussion

The topics to be covered are repeatability, deformation history effects, section size effects, comparison with long time creep data and applications of the SRT approach.

Repeatability. There are two major categories leading to variation in test results: precision and control of test variables, and variation in actual properties due to chemistry and microstructural differences. Many studies have shown results of conventional creep tests may vary appreciably between laboratories and even within laboratories due to calibration problems, temperature control, and heating method, specimen alignment, and loading procedure. Also the onset of fracture processes may complicate creep strength evaluation. No similar comparison is available for the SRT method, although it may be stated that the test requires the use of more sophisticated equipment and careful testing procedure. For the traditional creep test, both categories of scatter are normally lumped together to define a broad scatter band of creep data points. No attempt is normally made to determine the reason for a point to be on the lower bound of the scatter band. This necessitates the use of a safety factor on design stress which may lead to an overly conservative approach to creep design.

Although the role of test procedure in the SRT approach should not be dismissed, it is becoming increasingly clear that specimen to specimen differences may be accounted for by actual micro-

structural differences in the specimens. For example, the four highest strength specimens in Fig. 11 also had the highest elastic moduli. Also, in the SRT method, a wide range of creep rates is evaluated for each specimen rather than a single creep rate point as in the traditional approach. This suggests that a worst case microstructural condition may be evaluated for design. By recognizing also that the stress-creep rate curve is a locus of initial creep rates as a function of stress, the minimum creep rate will be at or lower than the SRT rate for the same condition. Thus, the SRT test has the potential to provide an absolute lower bound on creep strength and require no additional safety factor on stress.

Only one condition was repeated for ostensibly similar material. Specimens $7L$ and $7L$ rpt were taken from the same shank area and tested under the same conditions. However, the slightly higher strength in $7L$ was consistent with a slightly higher modulus (140,000 MPa versus 135,000 MPa). It could, therefore, reflect a genuine microstructural difference rather than a test variance.

Deformation History Effects. There are separate considerations of the effect of strain history and temperature history. These were not studied specifically in the current project. However, because of the crossover of the stress versus creep rate curves, it is likely that complex strain hardening and recovery processes would lead to significant strain history effects. Also, the presence of time-dependent anelasticity would contribute to these complexities. Generally, anelastic contributions to total creep strain become more significant at lower stresses.

It is important to recall that there is little memory of prior thermomechanical history for subsequent straining as indicated by the similar yield stress and flow curve in each successive test, especially in the larger specimens. This suggests that service deformation and thermal exposure may have only a small effect on creep strength as measured in the SRT test. It also implies that differences measured for the different locations most probably indicate initial differences rather than major changes occurring during service. This statement does not necessarily apply to the CDR results which are considered subsequently.

Section Size Effects. For the basic evaluation of creep strength, four standard specimens were taken from the blade shank. These were tested at 800C, 850C, and 900C with a repeat series at 900C. In addition, a miniature specimen ($4b$) was taken from the same location and tested at 850C. Two SRT runs from 0.4 percent and 0.8 percent strain were completed and a final CDR test was run at 800C. The SRT data provided a basis for evaluating possible section size effects on creep strength of IN738.

In fact Fig. 11 shows that for specimens $4b$ and $6L$ the creep strengths were very similar. The tensile data were also nearly identical (compare Figs. 1 and 9) and the same modulus of 140,000 MPa was measured.

Figure 11 indicates that for specimens taken from the same location, implicitly with a similar microstructure, there may be no significant effect of test section size. This is a particularly interesting and important observation. Previous studies have demonstrated a strong effect on rupture life of section size (reduced life in smaller sections) in equiaxed castings [8] but not in longitudinally oriented DS castings or monocrystals of the same alloy [9]. One likely explanation is that the reduced rupture life is associated with embrittlement due to oxygen penetration and environmentally enhanced intergranular cracking.

The effect is absent if there are no transverse grain boundaries. In the present work the results should not be influenced by fracture processes and they do in fact indicate that the creep strength in equiaxed castings is not affected by section size. Therefore, we anticipate that there would be section size effects on fracture resistance. Unfortunately, all our CDR tests have been on miniature specimens, so confirmation is not possible based on the present results.

These results on specimens $4b$ and $6L$ indicate that the creep strength of a particular microstructure (same grain size, dendrite spacing, and modulus) should be independent of test specimen size to a good approximation. However, the results reported in the bulk of this report indicate that the creep strength of different sections within a component may vary considerably. Because of this, we cannot evaluate unambiguously the effect of engine operation on creep strength. Evidence was presented to indicate that major thermomechanical excursions (i.e., prior SRT testing) may have little effect on creep strength, implying that modifications in precipitate distribution and dislocation structure are less important than the starting microstructural conditions identified above. It would clearly be of great value to do a comprehensive evaluation of a new blade.

Unfortunately, the CDR test on $4b$ broke outside the gage section, and the stress record showed an elastic unloading.

Comparison With Long Time Creep Data. Experience to date in comparing SRT results with minimum creep rates in a variety of alloys has generally indicated lower creep rates than the SRT results with the differences decreasing with increasing test temperature. It appears that at lower temperatures and higher stresses primary creep is significant in the creep tests. Conversely, at higher temperatures and lower stresses the minimum creep rate is likely to be very similar to the SRT curve. In the parametric comparison of Fig. 6, for example, *mcr* results for the same heat and grain size would tend to merge at the lower stresses. This is consistent with what was stated previously regarding the SRT data being capable of defining a lower bound behavior for creep strength.

In the creep rupture test, when the life of a specimen is short due to embrittlement or casting defects, measurements of rupture life and perhaps also times to specific strains will be influenced by fracture processes and hence comparison will have less meaning. This appears to have happened with the two creep test results.

Application of the SRT Approach. There are several ways in which the SRT data may be used in optimizing processing, as a basis for creep design, and for life management decisions. The construction of pseudo-stress-strain curves such as those of Fig. 4 may be used directly. For example, isochronous stress-strain curves constructed by cross plotting from conventional creep curves are often used in design of metals and plastics. In the same manner, the present curves may be used where time is estimated from strain divided by creep rate at any point on the curves or, preferably, directly where the parameter is creep rate.

Since all the curves are derived from the basic stress versus creep rate curves, it is recommended that these curves should form the basis of any design and analysis procedure. Moreover, creep rate may be used as a state variable whereas time requires a zero reference point and cannot be usefully employed as a state variable. Also, the SRT data for 0.4 percent strain involves minimal change in state during the relaxation run. Thus, it provides a true measure of the creep strength for a given state. In contrast, in the traditional long time creep testing the state is changing continuously as a function of strain and time.

A possible foundation for a new design methodology based on the SRT test is presented in terms of the parametric representation of Fig. 6. This master curve may be used, based on three one day tests, to compare processes, heat treatments, service exposures etc. It may also be used as a basis for creep design. As an example, possible design points at different temperatures are identified on the figure at creep rates of $3\times10^{-11}\,s^{-1}$. This is a rate corresponding to 1 percent in 100,000 h. However, calibration with current design practice might lead to the choice of a different creep rate.

Finally, we consider the combined use of the SRT and CDR tests for component life management decisions. For this alloy, and many other high strength superalloys, it appears that loss in creep strength defined from the SRT tests, as a result of thermal mechanical exposure, is likely to be minimal. However, embrittlement due to intergranular penetration of gaseous species, or due to

grain boundary segregation of harmful species from the matrix, may occur. This is expected to proceed progressively to a point where the fracture resistance is no longer acceptable for the component requirements. There are two issues here: at what level is the strain at failure unacceptable, and is the location of the embrittled region critical for the safe operation of the component? Suppose this unacceptable level is set at 2 percent. The critical areas are identified for each blade design, the level of embrittlement is measured periodically, and end of life is defined as the time when the lower limit is reached. On this criterion, if the locations for $1b$ and $2b$ are considered critical, i.e., local failure could lead to part failure, then the blade has reached the end of life. Engineering judgment must, of course be applied here. In a separate study with small solid IN738 blades, the trailing edges were found to be severely embrittled, but this was judged not to be a critical location for part failure.

One aspect of the problem of remaining life assessment of turbine blades which has not been addressed is the difficulty in knowing the extent to which the measurements reflect differences that were present when the components were first put in service. It has become common practice to make measurements on the thick section blade shanks to serve as the undamaged condition for comparison. However, there is increasing evidence suggesting that very significant differences exist initially between the properties of these thick sections and the thin wall airfoil sections. In cast blades the thicker section with lower solidification rates results in larger grain sizes and dendrite arm spacing. This in turn results in higher creep strength. The present data are consistent with this observation and no clear evidence for loss in creep strength due to service exposure could be obtained. It is worth noting that basic design data are normally measured on large cast to size specimens or on specimens taken from large cast blocks. There is no formal procedure to use properties unique to the local solidification conditions in analysing blade creep. The tensile ductility measurements at 800C do, nevertheless, suggest that appreciable embrittlement has occurred in thin sections near the pressure side of the blade airfoil. Clearly, resolution of these issues requires a systematic study of the properties of the material at different locations in new blades. The SRT and CDR techniques appear to be well suited to such a study.

The traditional approach for life prediction uses extrapolated time to rupture of a specimen as a direct indicator of remaining part life. However, the hybrid nature of the measurement, i.e., an arbitrary combination of creep strength and fracture resistance, leads to a fundamental conceptual flaw. This has been termed the ''Remaining Life Paradox.'' Material taken from a failed part will invariably have a finite life in a subsequent creep rupture test. Therefore, it follows, that life measured in a laboratory stress rupture test of a specimen taken from a part may have no bearing on the remaining life of the part. By decoupling the two properties and introducing separate creep strength and fracture resistance criteria, as outlined above, the paradox is avoided. End of life is then defined unambiguously in terms of minimum acceptable performance levels.

Conclusions

1 The stress relaxation test as conducted and analyzed here is a self-programmed variable stress test which can cover an enormous range of creep rates in a short time.

2 The data may also be presented in a form that is more consistent with current design practice, either in terms of pseudo stress versus strain as a function of creep rate or as stress versus time for a specific creep strain.

3 Repeatability of test data was excellent so that differences observed in different locations in a serviced blade and vane could be attributed to real difference in creep strength.

4 These differences most probably were due principally to initial differences in grain size and orientation rather than to microstructural evolution or damage developing during service.

5 A particular region of the blade was embrittled and this was attributed to the service exposure and most likely a result of oxygen penetration.

6 Limited long term creep data were quite comparable to the predictions using the SRT approach.

7 The stress versus creep rate plot derived from the SRT results may be parametrized and used as a basis for creep design and evaluation.

8 A life management procedure using a combination of SRT and CDR tests is outlined in which end of useful life is defined in terms of minimum performance levels.

Acknowledgment

This work was made possible by financial support provided by industry. Acknowledgement is also made of the fact that the reporting in this paper of the work was aided considerably by discussions with industry.

References

[1] Woodford, D. A., 1993, ''Test Methods for Accelerated Development, Design, and Life Assessment of High-Temperature Materials,'' Mater. Des., **14**, No. 4, pp. 231–242.
[2] Woodford, D. A., 1997, ''Creep Analysis of Directionally Solidified GTD111 Based on Stress Relaxation Testing,'' Mater. High Temp., **14**, No. 4, pp. 413–420.
[3] Woodford, D. A., and Iijima, K., 1997, ''Creep Strength Evaluation, Design, and Life Management of Cr-Mo-V Rotor Steel Using Stress Relaxation Testing,'' *Advances in Turbine Materials Design and Manufacturing*, Strang, A., Banks, W. M., Conroy, R. D., and Goulette, M. J., eds., Inst. Mater. pp. 613–624.
[4] Grwzwinski, G. G., and Woodford, D. A., 1995, ''Creep Analysis of Thermoplastics Using Stress Relaxation Data,'' Polym. Eng. Sci., **35**, No. 24, pp. 1931–1937.
[5] Reif, S. K., Amberge, K. J., and Woodford, D. A., 1995, ''Creep Design Analysis for a Thermoplastic from Stress Relaxation Measurements,'' Mater. Des., **16**, p. 15.
[6] Woodford, D. A., 1996, ''Creep Design Analysis of Silicon Nitride Using Stress Relaxation Data,'' Mater. Des., **17**, No. 3, pp. 127–132.
[7] Woodford, D. A., 1998, ''Stress Relaxation, Creep Recovery, and Newtonian Viscous flow in Silicon Nitride,'' J. Am. Ceram. Soc., **81**, No. 9, pp. 2327–2332.
[8] Oxx, G., 1973, ''Int. Conf. on Creep and Fatigue in Elevated Temperature Applications,'' Inst. Mech. Eng., Paper C212, Philadelphia, PA.
[9] Harrison, G. F., and Tilly, G. P., 1973, ''Int. Conf. on Creep and Fatigue in Elevated Temperature Applications,'' Inst. Mech. Eng., Paper C222, Philadelphia, PA.

Dae-Young Kim
Jong-Hyun Hwang
Kwang-Soo Kim
Joong-Geun Youn

Hyundai Industrial Research Inst.,
Hyundai Heavy Industries Co., Ltd.
1 Cheonha-dong Dong-ku Ulsan,
Korea 682-792

A Study on Fusion Repair Process for a Precipitation Hardened IN738 Ni-Based Superalloy

Several fusion repair processes such as laser cladding, laser welding and gas tungsten arc welding have been taken into consideration for repairing IN738 precipitation hardened Ni-based superalloy material. Effect of heat input on weld cracking susceptibility has been studied to obtain optimum condition for crack-free welds. Variations in cracking susceptibility as a function of welding heat input is discussed with reference to metallurgical characteristics of the welds. [S0742-4795(00)02003-2]

1 Introduction

Turbine blades in land based power generating combustion turbines burning low grade fuels frequently suffer from premature failures such as airfoil tip loss due to hot corrosion and low cycle fatigue cracking. Refurbishment of the damaged turbine airfoil is commonly implemented due to its low cost as compared to replacement by new parts. Several repair processes such as welding, brazing and spray deposition have been introduced and are well-established for actual commercial repair applications [1]. Among these repair processes fusion welding process is very competitive one due to a relatively low initial investment. However, it is well known that high strength precipitation hardened nickel based superalloys possess poor weldability, for example cracking in the weldment during welding and post weld heat treatment [2]. This cracking tendency is known to increase with an increase in aluminum and titanium content, which are γ' precipitate former. Specialized techniques such as high temperature preheating and employment of lower strength filler metals have been tried to prevent such cracking problems [3]. These difficulties in controlling weld cracking have resulted in repair limitations to the hot gas path components, especially turbine blades. If it is possible to enhance repair area for the high strength precipitation hardened nickel based superalloy, it will contribute substantially to the reduction of life cycle cost of the combustion turbine.

Effect of welding heat input on cracking sensitivity of the IN738, a precipitation hardened nickel base superalloy, using gas tungsten arc(GTA) welding and laser welding has been studied to obtain optimum condition for crack free welds of IN738 alloy. Optimization of laser cladding process for IN738 alloy was also conducted to look into the possibility of introducing it as an alternative process for replacing preheating arc welding process for the alloy.

2 Experimental Procedure

Bead-on-plate welded specimens were made using cast IN738 plate and filler metal alloy. Each plate specimen was 25 mm×200 mm×5 mm in size. Chemical compositions of IN738 base metal and IN738 filler metal powder used for laser cladding studies are shown in Table 1. Gas tungsten arc welding (GTAW) conditions and laser welding (LW) conditions are summarized in Table 2 and Table 3, respectively. Dye penetrant inspection was conducted to evaluate the propensity of cracking after welding. Microstructural evaluation and hardness tests were also conducted to quantify the differences in cracking susceptibility of IN738 welds as a function of welding heat input.

3 Results and Discussions

Effect of Welding Heat Input—Gas Tungsten Arc Welding. Cracking rate of the autogenous GTA welds in IN738 alloy decreased as the heat input decreased from 9.9 to 5.5 kJ/cm as shown in Fig. 1. The data suggest a heat input under 5.5 kJ/cm is required for producing crack-free welds. Cracks were observed both at the fusion line and at the weld metal itself along dendrite boundaries as shown in Fig. 2. Fig. 3 showed the effect of heat input on the solidification microstructure of the IN738 GTA weld. As heat input increased, morphologies of the weld metal changed from fine columnar to coarse cellular shape. Dendrite arm spacing decreased as welding heat input decreased as shown in Fig. 4. It could be concluded that finer dendrite microstructure could reduce cracking susceptibility of the IN738 weldment by decreasing effective stress acting along the dendrite boundaries.

Effect of Welding Heat Input—Laser Welding. No crack was observed in the autogenous laser weldment of IN738 alloy in the heat input range of 2 to 8 kJ/cm. Microstructure of the weld reveals finer dendrite microstructure, compared with the GTA weld, Fig. 5. The dendrite arm spacings shown in Fig. 6 lie in the range of 6 to 10 μm indicating much finer spacing than that observed in GTA weld (20 to 50 μm). These finer microstructures in the weld metal lead to an improvement in the cracking resistance of IN738 laser weldment. Laser weldment showed HAZ softening depending on the heat input applied as shown in hardness distribution in Fig. 7. Even though this HAZ softening behavior impairs the HAZ cracking resistance of the precipitation hardened nickel based superalloy [4], this study show that it is possible to obtain crack free weldment by laser welding due to a finer weld microstructure. This indicates that laser welding of IN738 alloy enhances applicable heat input range without any cracking problems due to a relatively faster cooling rate than the GTA welding.

Laser Cladding. Laser cladding of the IN738 alloy using matching IN738 powder filler metal was conducted. There was no occurrence of weld cracking in the as-welded condition as shown in cross sectional microstructure of Fig. 8. Based on this study, laser cladding can also be used as an alternative fusion process to

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-313. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

Table 1 Chemical compositions of IN738 base metal and filler metal powder

Descriptions	Chemical Composition, wt.%							
	Ni	Cr	Co	Al	Ti	W	Mo	Ta
Cast IN738	Bal.	15.9	8.34	3.42	3.32	2.76	1.68	1.76
IN738 Filler Metal Powder	Bal.	15.9	8.34	3.25	3.37	2.77	1.76	0.83

Table 2 Process variables of gas tungsten arc welding employed in this study

Descriptions	Process Variables			Heat Input*1), kJ/cm
	Current, A	Voltage, V	Speed, mm/min	
GTAW-5.5	138	12	180	5.5
GTAW-7.5			130	7.6
GTAW-9.5			100	9.9

*1) Heat Input(kJ/cm) = (Current(A) X Voltage(V) X 60 X 10) / (Speed(mm/min.) X 1000)

Table 3 Process variables of laser welding employed in this study

Descriptions	Process Variables			Heat Input*1), kJ/cm
	Power, kW	Working Distance, mm	Speed, mm/min	
LW-2.0	2.0	220	600	2.0
LW-4.0			300	4.0
LW-6.0			200	6.0
LW-8.0			150	8.0

*1) Heat Input(kJ/cm) = (Power(kW) X 60 X 10) / Speed(mm/min.)

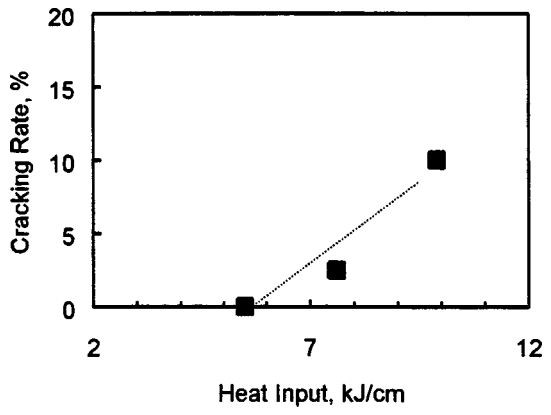


Fig. 1 Cracking rate in IN738 GTA weldment as a function of welding heat input (Cracking Rate=(Crack Length/Total Weld Length) ×100, percent)

conventional GTA welding. Hardness distribution of the laser clad layer also indicated the softening behavior in heat affected zone as shown in Fig. 9.

4 Conclusions

The following conclusions could be drawn from this study with regard to the effect of welding heat input on cracking susceptibility in IN738 weldments:

1 the lower heat input in GTA welding of IN738 alloy results in crack-free weldments by forming fine dendrite microstructure
2 laser welding of IN738 alloy enhances applicable heat input range without any cracking problems due to a relatively faster cooling rate than the GTA welding
3 laser cladding process can be chosen as an alternative repair process to GTA welding with high temperature preheating

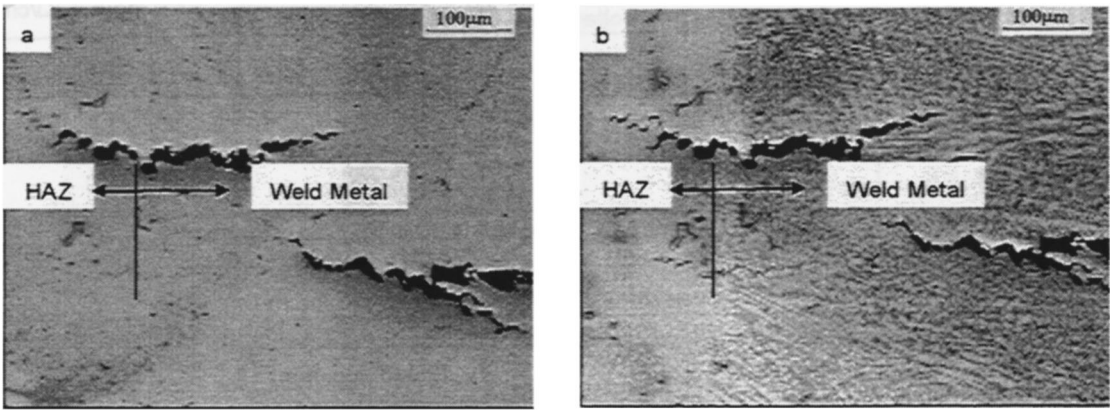


Fig. 2 Features of cracks observed in the IN738 GTA weldment; (*a*) before and (*b*) after etching

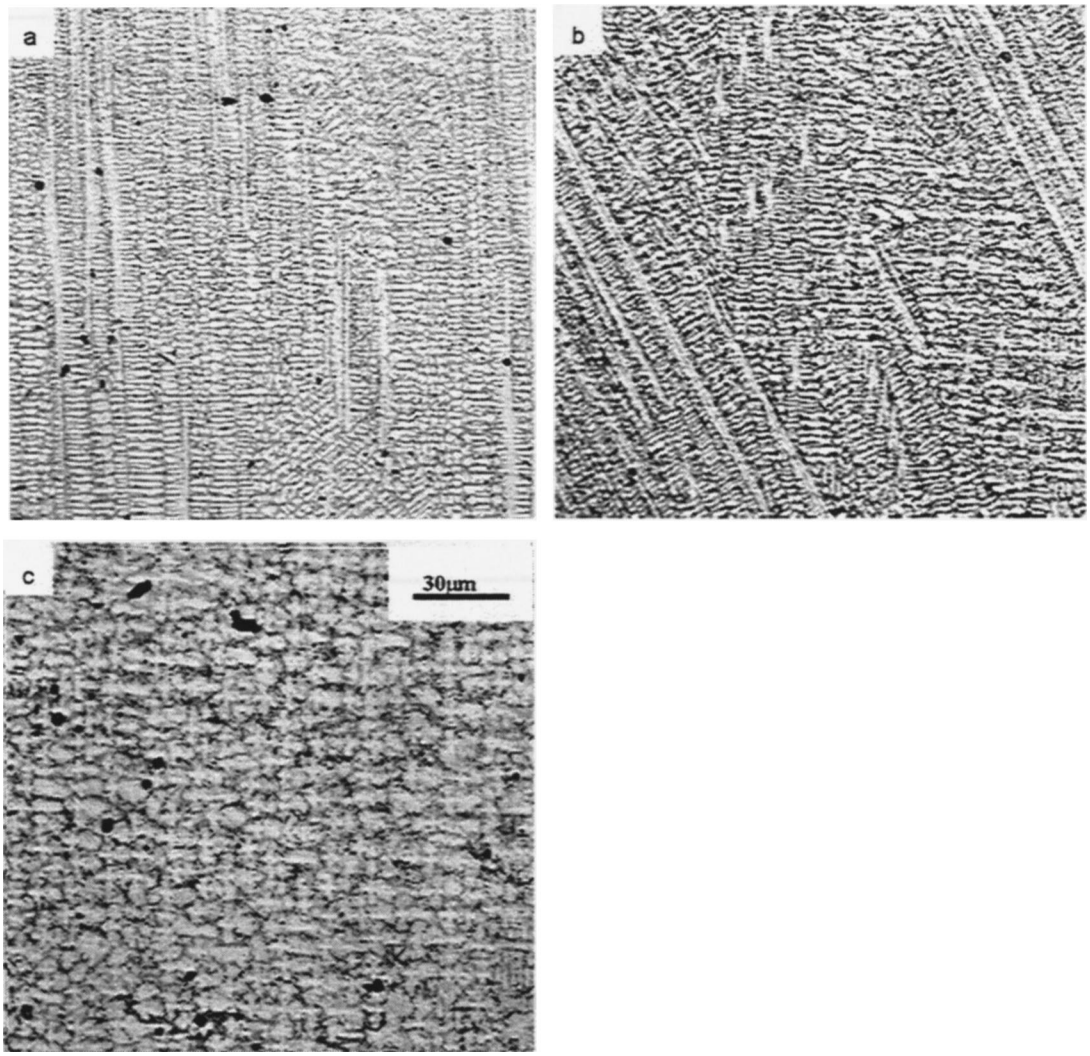


Fig. 3 Microstructures of IN738 GTA weld near fusion line as a function of welding heat input: (*a*) 5.5 kJ/cm; (*b*) 7.6 kJ/cm; and (*c*) 9.9 kJ/cm

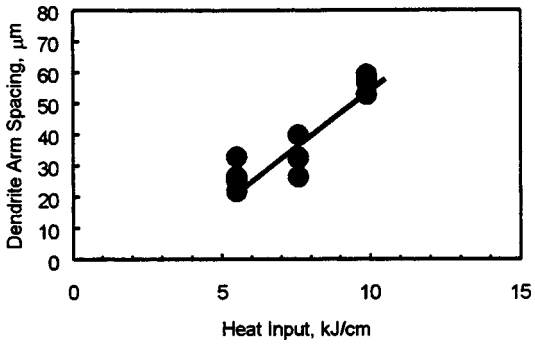


Fig. 4 Dendrite arm spacing in IN738 GTA weld as a function of welding heat input

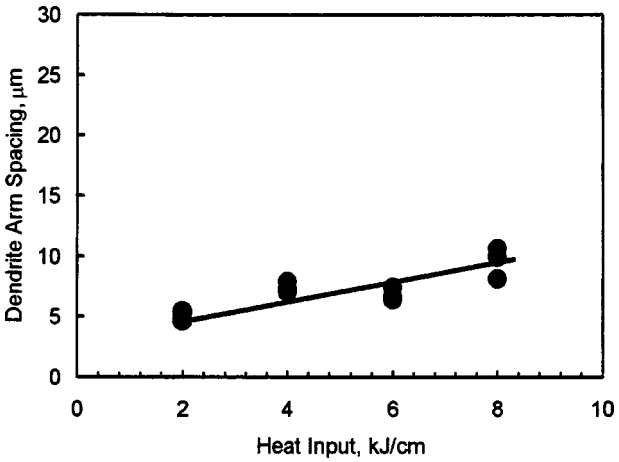


Fig. 6 Dendrite arm spacing in IN738 laser weld as a function of welding heat input

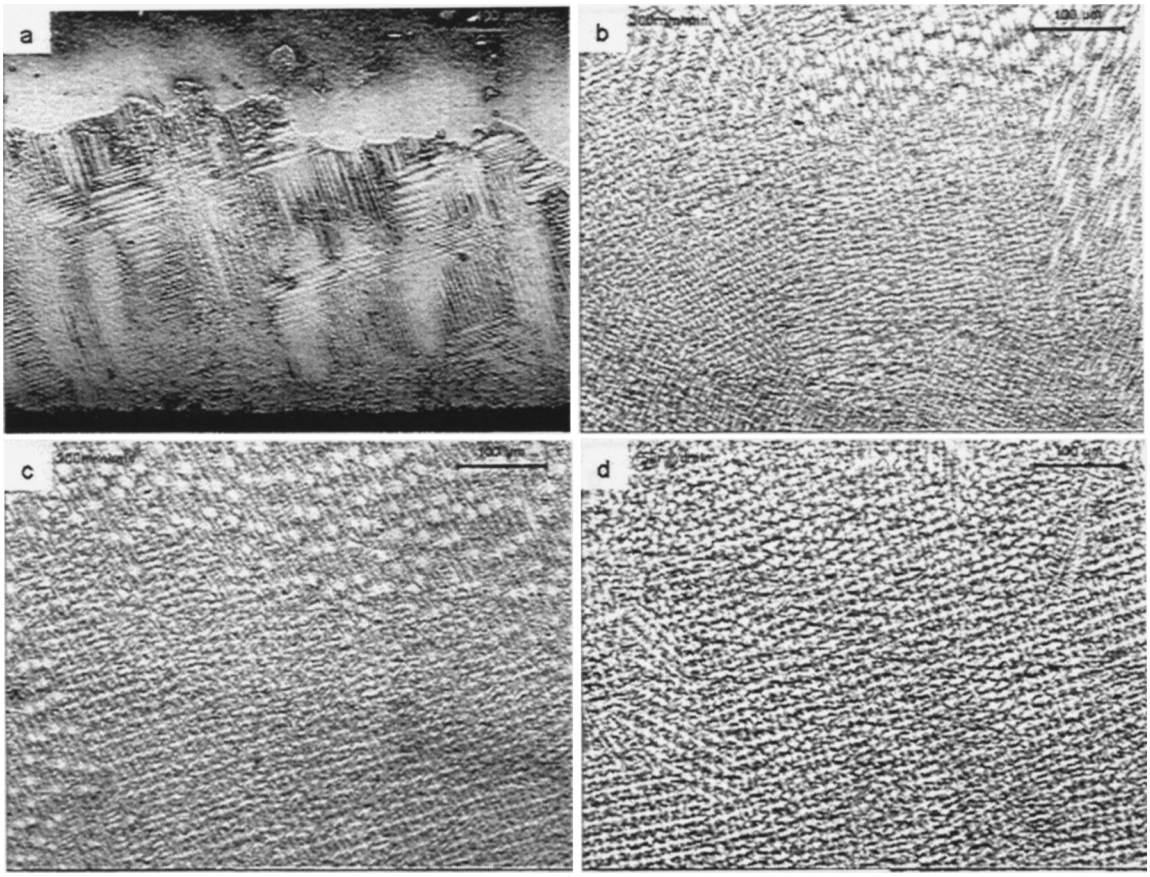


Fig. 5 Microstructures of IN738 laser weld near fusion line as function of welding heat input: (*a*) 2 kJ/cm; (*b*) 4 kJ/cm; (*c*) 6 kJ/cm; and (*d*) 8 kJ/cm.

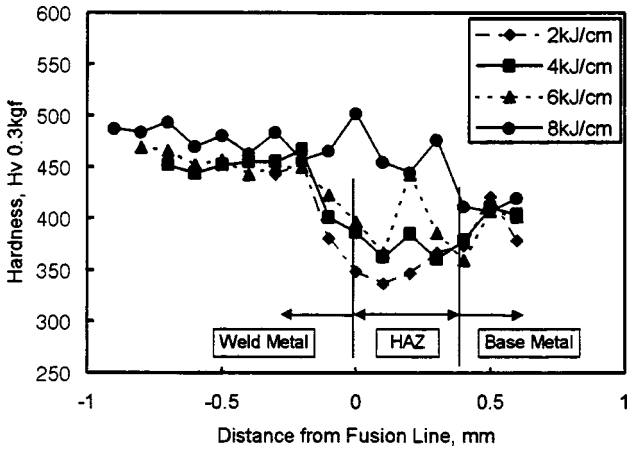


Fig. 7 Hardness distribution in laser weldment as a function of welding heat input

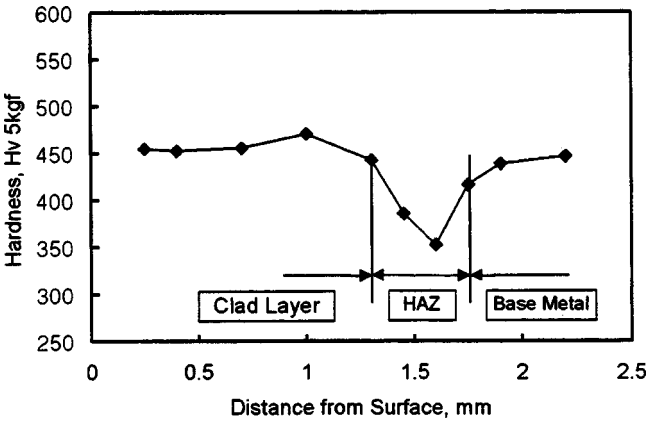


Fig. 9 Hardness distribution in the laser cladding layer

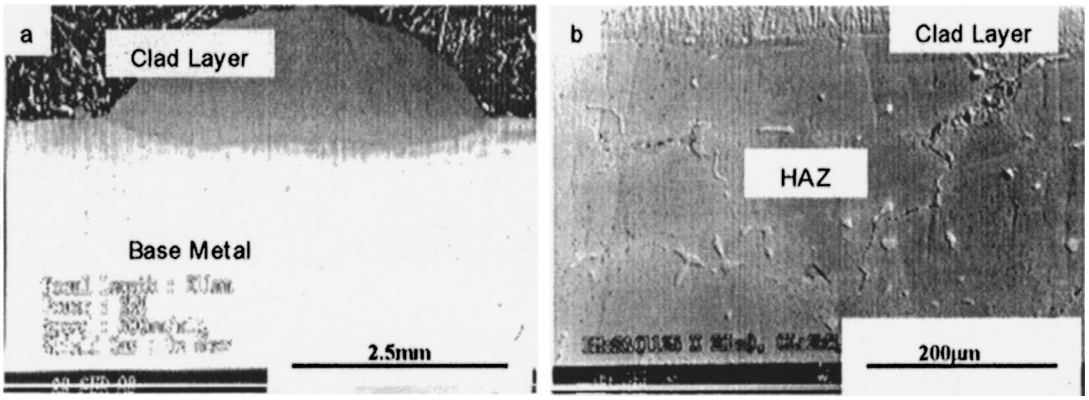


Fig. 8 Cross-sectional microstructures of laser clad layer: (*a*) lower magnification; (*b*) higher magnification.

References

[1] Dodgson, J. S., 1982, ''Repair Techniques for Hot Gas Path Components in Industrial Gas Turbine,'' ASME Paper 82-GT-104.

[2] Liburdi, J., and Lowden, P., 1985, ''Repair Techniques for Gas Turbine Components,'' AGARD Conf. Proc., No. 398, Sept., **22-1**.

[3] Jahnke, B., 1982, ''High-Temperature Electron Beam Welding of the Nickel-Base Superalloy IN-738LC,'' Weld. J. (London), **61**, No. 11, p. 343.

[4] Haafkens, M. H., and Matthey, J. H. G., 1982, ''A New Approach to the Weldability of Nickel-Base As-Cast and Power Metallurgy Superalloys,'' Weld. J. (London), **61**, No. 11, p. 25.

Anthony J. Smalley

David A. Mauney

Southwest Research Institute,
San Antonio, TX 78228-0510

Daniel I. Ash
Conoco Inc.,
Ponca City, OK 74602-1267

Sam L. Clowney
Tennessee Gas Pipeline Company,
Houston, TX 77252-2511

George P. Pappas
Union Gas Limited,
Chatham, Ontario,
Canada N7M 5M1

Evaluation and Application of Data Sources for Assessing Operating Costs for Mechanical Drive Gas Turbines in Pipeline Service

This paper evaluates and demonstrates how the public domain data provided by individual interstate pipeline companies to FERC, when combined with individual company equipment lists, can be used to regress industry information on cost of operations and maintenance, fuel gas used, and cost of fuel and power. The paper describes the methods of analysis and identifies their limitations. The paper presents results of such regression analysis as average and variance of cost and fuel usage for industrial gas turbines and aeroderivative gas turbines. It provides further comparisons between gas turbine prime movers, reciprocating engine prime movers, and electric motor drives, and presents annual costs per installed horsepower as a function of turbine size. The paper is based on work performed for PRC International and the Gas Research Institute.
[S0742-4795(00)01003-6]

Introduction

Mechanical drives for centrifugal compressors provided an early application for land based gas turbine engines. Today the gas transmission industry has deployed both industrial and aeroderivative gas turbines widely. Reciprocating integral engine compressors still dominate the U.S.A. installed base, but many companies now prefer a gas turbine driven centrifugal compressor for new installations. In 1994, gas turbines represented about 20 percent by number of the U.S. installed gas transmission fleet, and 30 percent by horsepower. In the rest of the world, including Canada, gas turbines dominate the fleet more heavily with over 90 percent of the installed horsepower. As a recent trend worthy of note, pipeline companies in both the U.S. and Canada have chosen to install motor driven compressors as a significant fraction of their new or replacement horsepower.

The U.S. gas transmission industry continues to deregulate and faces increasingly aggressive competition. Reducing the cost of gas operations offers a way to compete more effectively, and to maintain profitability. Knowledge of current costs can help a company control these costs especially when expressed as an intensity (which relates cost to business activity). However, the need for individuals to document and gather specific cost data conflicts directly with increasing demands for individual productivity.

The following results show how public domain information, combined with the most readily available company information, can help profile the industry's compressor operations and maintenance costs; it further uses linear regression to infer cost intensities by type of driver (even without cost data recorded to this level of detail). The results of the paper emphasize gas turbine costs, but also provide data for comparison between gas turbine, reciprocating engine, and electric motor drives. Operating costs addressed include compressor station maintenance, compressor operation, and fuel usage. The paper presents results with particular relevance to gas turbine operation from a detailed investigation of compressor operating costs by Smalley et al. [1].

Contributed by the International Gas Turbine Institute (IGTI) of The American Society of Mechanical Engineers for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-51. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

Data Sources

The Federal Energy Regulatory Commission (FERC) requires all U.S. interstate pipelines to file detailed operating statements each year. This data includes the following details of compressor operations by pipeline company, for each calendar year:

- pipeline total compressor station maintenance costs for gas transmission
- pipeline total gas deliveries
- station-by-station fuel use
- station-by-station O&M costs
- station-by-station power and fuel costs

Perseverance will lead to success in downloading this data! It provides a basis for normalizing corporate performance (e.g., costs and fuel burned per unit volume delivered); however, it provides no direct basis for normalizing cost and fuel use relative to work done (horsepower hours) or installed power; this lack of measures for engine size and operation could limit the basis for assessing and controlling costs.

Fortunately, however, most operating companies have, in readily transferable electronic form, a detailed definition of their installed horsepower, comprising at each compressor station:

- unit number or identifier
- driver type
- model number
- installed horsepower
- year installed

Assuming cost or consumption by station varies linearly with installed horsepower, linear regression provides the constants of proportionality, and measures of variance. By combining data for all companies, we can infer the costs and fuel use per horsepower for each compressor type and model which has sufficient representation in the database.

The FERC data base does not contain maintenance cost alone at the station level (just at the company total level); thus, while most companies keep records of this quantity by station for rolling up

into the company total, acquiring data on maintenance cost by station requires contact with each individual company, and a request to provide this detail.

Further review shows that the FERC data base does not include horsepower hours; thus establishing a measure of compression work done at the company, station, or unit level (whichever might be available) also requires direct contact. In summary then, the following information was obtained by request from individual companies:

Company Horsepower Totals:	28 Companies	17 Million Horsepower
Horsepower Details at Each Transmission Compressor Station:	20 Companies	14 Million Horsepower
Total Annual Horsepower Hours:	15 Companies	10 Million Horsepower
Horsepower Hours by Compressor Station:	5 Companies	8 Million Horsepower

Methods of Analysis

Linear regression provides the primary analytical technique used to extract information from the data base, yielding constants of proportionality or slopes, standard deviations, and standard errors. The company level analysis profiles the industry with regressed constants of proportionality linking measures of company cost to measures of company size or activity. It further helps assess consistency in relationships between these measures as a guide to their application for station level analysis.

In the company level analyses, each pipeline company provides pairs of dependent and independent variables. For example, to regress compressor station maintenance cost against deliveries requires extracting from the data base values for each company's maintenance cost and deliveries over the year in question. Because the completeness of the database varies slightly for the different quantities, a typical regression involves 26 to 30 companies, and in the case of horsepower hour regressions, it only involves about 15 companies.

The station level analysis extracts information below the level at which most companies keep records—by inferring dependence of cost and consumption on prime mover type, and on individual model.

The station level analyses involve a larger, more complete database of information than the company analyses. In general, the station level analysis uses multiple linear regression involving significantly more than one independent variable. The simplest station level analysis seeks constants of proportionality between cost and horsepower for six prime mover types:

- two-stroke reciprocating engine
- four-stroke reciprocating engine
- separable reciprocating engine
- electric motor
- aeroderivative gas turbine
- industrial gas turbine

Every station with available cost information comprises one or more of these engine types, quantified by the installed horsepower for that type. The analysis assumes cost varies linearly with horsepower, and extracts constants of proportionality, which minimize the RMS deviation between station cost and predictions using these constants. The database contains a large number of compressor stations (many hundreds). Many stations comprise only one or two types and for more sparsely represented types, it was found beneficial to regress the constants of proportionality from a series of subsets of the data base, each of which contained only stations with the particular type of interest; thus, in particular, we regressed electric motor information only from the reduced set of stations which contain electric motor drives. Similar station level regression analysis was performed to extract maintenance costs for individual models, and to infer maintenance costs as a function of size within a particular type.

As will be seen, all regression analyses produced results with uncertainty, subject to limitations imposed by the implied assumption that cost depends linearly on horsepower installed. A more desirable regression would be to seek size, type, and model information by horsepower hour regressions, but unfortunately, the available data did not support this.

Results

Figure 1 shows how a company's combined operations and maintenance cost for 1994 (on the vertical axis) correlates with deliveries for that year (on the horizontal axis). Each point on this figure corresponds to an individual U.S. interstate pipeline company; the line represents the regressed line which ''best fits'' the data by minimizing the root mean square total of deviations from that line. The quantity R^2 measures how strongly the dependent variable depends on the independent variable—$R^2=100$ percent would imply ''perfect dependence'' with no scatter; $R^2=0$ would imply no dependence. The value of 82.44 percent for R^2 in Fig. 1 indicates relatively strong dependence. The slope ($16.22/MMCF) represents the constant of proportionality between O&M cost and deliveries. The standard error ($6.55 million) measures the RMS deviation of the data points from the line, and provides a different measure of scatter.

Figure 2 shows how total fuel burned in gas transmission cor-

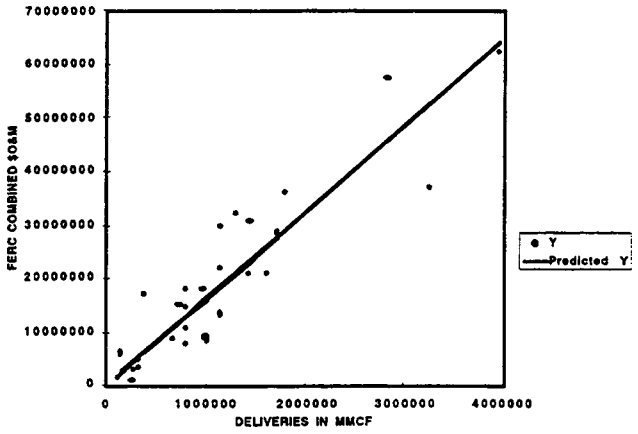


Fig. 1 Correlation of FERC combined $O&M with deliveries in MMCF–FERC filing U.S. interstate pipelines only—R^2 = 82.44 percent; std. error=$6.55 MM; slope=$16.22/MMCF.

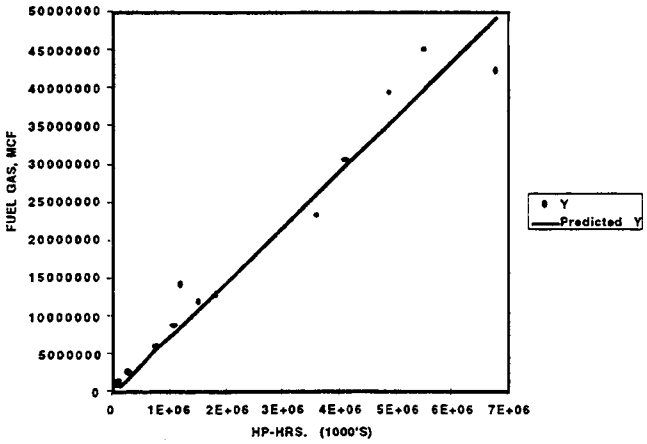


Fig. 2 Correlation of fuel gas (MCF) with HP-hrs. for total population with data—single largest outlier eliminated—R^2 = 96.05 percent; std. error=3.19 BCF; slope=7.25 MCF/1000 HP-hr.

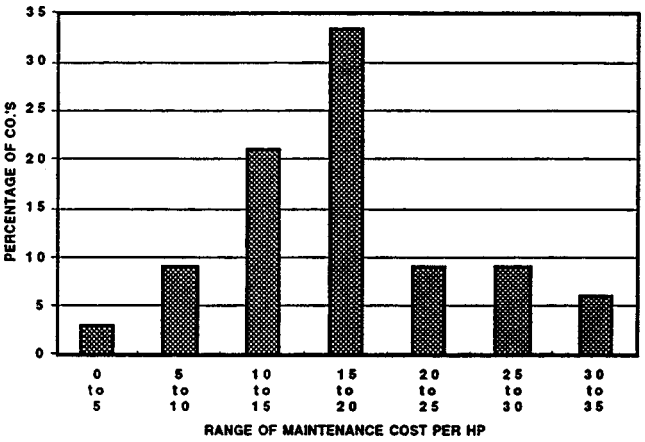


Fig. 3 Histogram for CSM cost per HP, showing percentage of Co.'s in each indicated range—industry mean=16.65; median =17.61.

relates with total horsepower hours. Here the R^2 has a value of 96.05 percent (for a set which includes nine U.S. interstate companies; three Canadian companies; and two U.S. intrastate companies). This indicates strong dependence, with a standard error of 3.19 billion cubic feet, and a slope of 7.25 MCF per thousand horsepower hours. Both Figs. 1 and 2 represent industry-wide profiles for gas transmission, undifferentiated by engine type, and comprising inevitably a mix of reciprocating and turbine engines.

Figure 3 shows an alternative way to profile the industry via a histogram. Here the bars indicate on the vertical axis the fraction (percent) of all companies considered which fall into each range on the horizontal axis. It indicates that the largest number (33 percent) spent between $15 and $20 per installed horsepower on maintenance; with a range from between zero and $5 on the low end, and $30 to $35 per horsepower on the upper end; combining the largest two bars, well over 50 percent of companies spend between $10 and $20 per installed horsepower on maintenance. Again, this histogram profile does not differentiate engine type.

Table 1 presents a broader summary of key correlation results for compressor station maintenance costs; it covers regression against deliveries (MMCF), installed horsepower, and 1000s of horsepower hours; it distinguishes between the entire population analyzed (which includes several million Canadian horsepower—25 percent of the total, and two intrastate pipelines), and FERC filing interstates. It shows values for correlation coefficient (R^2), standard error, slope, standard error in slope, lower and upper limits for 95 percent of the population (based on Gaussian analysis), and size of the population analyzed. In general, the FERC companies alone, and total population provide similar values for slope, even though the variance may differ significantly.

Table 1 Compressor station maintenance—summary of regression parameters

Regression	R^2, %	Std. Error	Slope	Std. Error in Slope	Std. Slope Error, %	Lower 95% Limit on Slope	Upper 95% Limit on Slope	Size
$CSM vs. MMCF	43.81	5811613	6.20	0.63	10.21	4.91	7.50	31
$CSM vs. HP	79.39	3504008	14.87	0.84	5.64	13.16	16.57	33
$CSM vs. 1000's of HP-Hrs.	58.90	4479962	3.33	0.37	11.11	2.53	4.13	16
$CSM vs. MMCF (f0)(1)	62.00	3782174	6.26	0.48	7.75	5.26	7.26	27
$CSM vs. HP (f0)	62.84	3740344	14.86	1.14	7.65	12.52	17.20	27
$CSM vs. 1000's of HP-Hrs. (f0)	71.20	3809422	3.97	0.39	9.77	3.13	4.82	13
$CSM vs. MMCF (1 outlier removed)	52.95	4126503	5.59	0.46	8.23	4.64	6.53	30

(1) f0 = FERC filing interstate pipelines only.

Table 2 Industry normalized averages, regression slopes, and medians

Regression Slopes, and Medians

Normalization	Population	Industry Average	Regressed Slope	Median Ratio
Maint. Cost vs. Deliveries	Total	$6.90/ MMCF	$5.59/ MMCF	$6.31/ MMCF
Maint. Cost vs. HP	Total	$16.65/ HP	$14.87/ HP	$17.61/ HP
Maint. Cost vs. 1000's of HP-Hrs.	Total	$3.98/ 1000 HP-Hr.	$3.33/ 1000 HP-Hr.	$5.26/ 1000 HP-Hr.
Maint. Cost vs. Deliveries	FERC Only	$6.94/ MMCF	$6.26/ MMCF	$6.42/ MMCF
Maint. Cost vs. HP	FERC Only	$17.43/ HP	$14.86/ HP	$17.61/ HP
Maint. Cost vs. 1000's of HP-Hr.	FERC Only	$4.44/ 1000 HP-Hr.	$3.97/ 1000 HP-Hr.	$5.25/ 1000 HP-Hr.
Combined O&M vs. Deliveries	FERC Only	16.64/ MMCF	$16.22/ MMCF	$15.22/ MMCF
Combined O&M vs. HP	FERC Only	$44.66/ HP	$38.64/ HP	$44.96/ HP
Combined O&M vs. 1000's of HP-Hrs.	FERC Only	$11.28/ 1000 HP-Hrs.	$10.92/ 1000 HP-Hrs.	$12.63/ 1000 HP-Hrs.
Fuel Gas vs. Deliveries	Total	1.42%	1.15%	1.06%
Fuel Gas vs. HP	Total	34.14 MCF/ HP	30.39 MCF/ HP	31.70 MCF/ HP
Fuel Gas vs. 1000's of HP-Hrs.	Total	8.25 MCF/ 1000 HP-Hr.	7.25 MCF/ 1000 HP-Hr.	7.92 MCF/ 1000 HP-Hr.
Fuel/Power Cost vs. HP	FERC Only	$62.23/ HP	$62.75/ HP	$58.5/ HP

Table 2 presents industry normalized averages, regression slope, and median ratios for maintenance cost, O&M cost, and fuel use versus deliveries, HP, and 1000s of HP-hours. An industry average equals the total industry cost divided by the total industry value of the normalizing quantity. The regression slope has previously been defined as the constant of proportionality, which best fits the individual company data points. The median ratio requires calculating the normalized cost for each company, then finding the median value for this normalized cost. While the values in each row of Table 2 have the same units, they indicate that significant variation exists as a result of differences in the method of calculation; thus the data should be applied with careful consideration of the quantity and its definition, and awareness of this variability.

Figure 4 presents results of compressor station regression

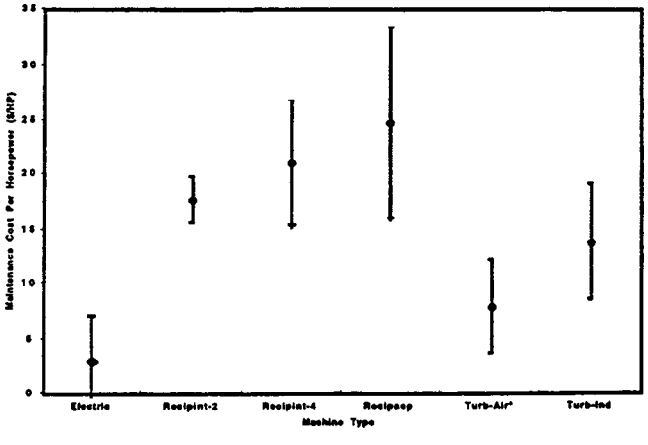


Fig. 4 Maintenance cost per HP as a function of machine type

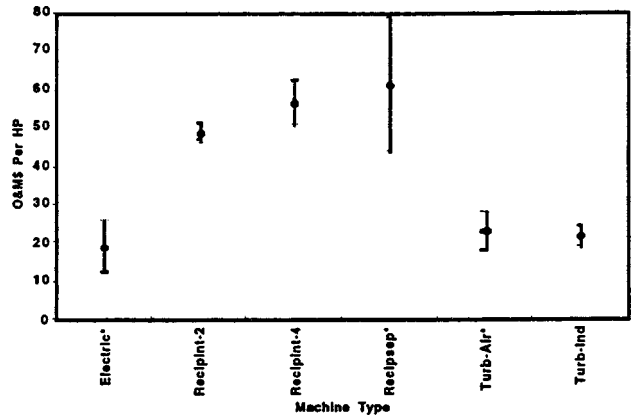


Fig. 5 O&M cost per HP as a function of machine type

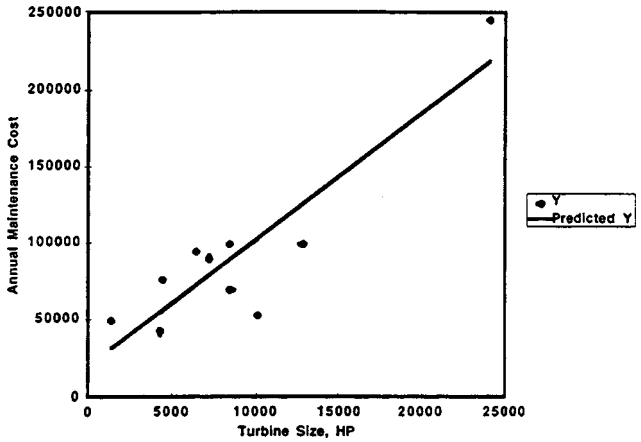


Fig. 6 Annual maintenance cost versus industrial gas turbine size, HP

analysis; it shows how inferred maintenance cost per horsepower varies with driver type, and includes error bars which indicate the range of mean ± one standard deviation for

- electric motor
- two-stroke reciprocating engine, integral with the compressor
- four-stroke reciprocating engine, integral with the compressor
- separable engine (not integral with the compressor)
- aeroderivative gas turbine
- industrial gas turbine

Figure 5 presents a similar comparison by type of drivers—this time for combined operations and maintenance cost per horsepower (without fuel cost). Comparison with Fig. 4 for maintenance alone indicates that combined cost of O&M is approximately 2.5 times the cost of maintenance alone; it also shows the error bands are a substantially smaller fraction of the average than the error bands for maintenance cost alone. In other words, the combined O&M cost produced a much tighter regression that maintenance alone.

It is possible to extract values for various costs as a function of model, as Smalley et al. [1] demonstrates. This paper seeks to demonstrate methods and data sources, rather than compare products, and presents these results as mean maintenance cost per year as a function of size for industrial turbine models. Each data point in Fig. 6 corresponds to a model; the size ranges from just over 1000 HP to about 25,000 HP. The figure includes a regression line. This figure, while exhibiting some scatter, suggests that the variation of cost with size has a non-zero intercept—that is a cost per turbine which does not depend on size. Linear regression quantifies the ''per turbine'' maintenance cost as $20,136 with a slope of $8.27/HP, and a standard error of $27,072; while this standard error confirms the scatter in the data, the general trends of the regression are as expected.

Discussion

The illustrative analyses based on the data from FERC, and data from individual companies indicate that much can be learned about the gas transmissions cost of compression, including turbine drives, both in comparative and absolute terms. Linear regression techniques can provide information to a level below that of record keeping practices. Information can be extracted on cost of maintenance, cost of O&M combined, and fuel use. Smalley et al. (1997) provides a more comprehensive presentation of possible results and methods of analysis.

Conclusions

Useful information regarding gas transmission maintenance can be extracted from public domain data in combination with supplementary information provided by individual companies.

1994 industry average values available include $16.65/HP for maintenance cost; $3.98/1000 HP-hours for maintenance cost; 8.25 MCF per 1000 HP-hours for fuel gas burned, and 1.42 percent for fuel gas volume per volume delivered. The methods discussed in this paper can be applied with benefit to more recent years of operation.

For aeroderivative gas turbines regressed values for 1994 were $7.83 and $22.81 per HP for maintenance alone and for combined O&M cost, respectively. For industrial gas turbines, these regressed costs for maintenance alone and for combined O&M were $13.75 and 21.50, respectively.

When regressed for individual models, the average maintenance cost for industrial gas turbines has the form

$$\$\text{Maint.} = \$20{,}136 + \$8.27/\text{HP}.$$

Acknowledgments

The authors wish to express appreciation to PRC International and the Gas Research Institute for their support of this work, and for permission to publish this paper. The authors also greatly appreciate all the help provided by individual pipeline companies in providing data for this investigation.

Reference

[1] Smalley, A. J., Mauney, D. A., and Ash, D. I., 1997, ''Compressor Station Maintenance Cost Analysis,'' final report PRCI No. PR-15-9529, Compressor Research Supervisory Committee of PRC International.

Actively Controlled Bearing Dampers for Aircraft Engine Applications

John M. Vance
Mechanical Engineering Department,
Texas A & M University,
College Station, TX 77843

Daniel Ying
Turbine Manufacturing Division,
ABB Power Generation, Inc.

Jorgen L. Nikolajsen
Staffordshire University,
POB 333,
Stafford ST18 ODF, U.K.

This paper describes some of the requirements for bearing dampers to be used in an aircraft engine and briefly discusses the pros and cons of various types of dampers that were considered as candidates for active control in aircraft engines. A disk type of electrorheological (ER) damper was chosen for further study and testing. The paper explains how and why the choice was made. For evaluating potential applications to aircraft engines, an experimental development engine (XTE-45) was used as an example for this study. Like most real aircraft engines, the XTE-45 ran through more than one critical speed in its operating speed range. There are some speeds where damping is desirable and other speeds where it is not. Thus, the concept of a damper with controllable forces appears attractive. The desired equivalent viscous damping at the critical speeds along with the available size envelope were two of the major criteria used for comparing the dampers. Most previous investigators have considered the ER damper to produce a purely Coulomb type of damping force and this was the assumption used by the present authors in this study. It is shown in a companion paper, however, that a purely Coulomb type of friction cannot restrain the peak vibration amplitudes at rotordynamic critical speeds and that the equivalent viscous damping for rotordynamics is different from the value derived by previous investigators for planar vibration. The type of control scheme required and its effectiveness was another criterion used for comparing the dampers in this paper. [S0742-4795(00)00803-6]

Introduction

A disk type of electroviscous damper, utilizing fluid with electrorheological (ER) properties, was invented in 1988 by J. Nikolajsen [1]. Figure 1 shows a cross section of the damper, with six thin nonrotating disks moving with the outer race of a ball bearing and with five nonrotating disks attached to the housing and sandwiched in between. Since the ball bearing is on a squirrel cage (flexible) support, transverse rotor motion at the bearing shears the fluid between the disks. Nikolajsen and Hoque [2] constructed a rotordynamic test rig and demonstrated experimentally that the damper could substantially reduce the peak rotor vibration amplitude at a low critical speed (1400 rpm). They also demonstrated that varying the voltage applied across the disks could control the effective damping. They hypothesized that active control could eliminate the critical speed response by changing the effective system *stiffness* (not the damping) at appropriate speeds during runup of the rotor. With regard to the damping, they state that ''the electroviscous fluid provides Coulomb-type friction damping at a variable rate controlled by a DC voltage applied to the fluid.''

It has been known for about a century [3] that electrorheological (ER) fluids can be concocted with shear properties that are controllable, by using an imposed electric voltage. Nikolajsen and Hoque [2] give some of the history of their development and give a useful description of one class of these ER fluids as shown in Fig. 2. In such fluids, the applied voltage converts them from a Newtonian fluid to a Bingham plastic, where the viscosity is not changed but a positive shear stress must be applied before the fluid breaks loose and begins to shear. A high enough voltage makes the fluid appear to be solid.

Shortly after Nikolajsen's invention, a study was carried out to determine the feasibility of developing actively controlled bearing dampers for aircraft engines. The ER damper was one of several candidate devices to be considered. This paper describes some of the requirements for bearing dampers to be used in an aircraft engine and briefly describes the pros and cons of the various dampers that were considered for active control. Vance and San Andres [4] explain how active control could be implemented with the ER damper if the Coulomb model is a correct representation.

Vance and Ying [5], describe experiments with the actively controlled ER damper, carried out after this study was completed, that explored some of the practical rotordynamic considerations by conducting higher speed tests and implementing two different control schemes. Their experiments revealed the complex nature of the damping forces obtained from an ER bearing damper. The results of the experiments were not particularly encouraging in a

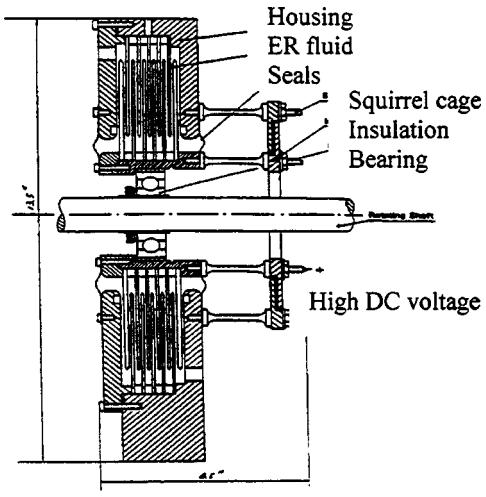


Fig. 1 Cross section of the ER damper

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper No. 99-GT-18. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

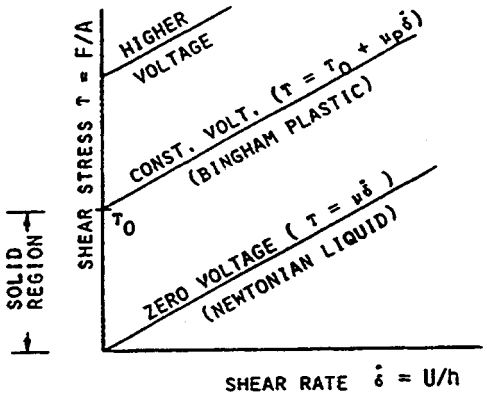


Fig. 2 Electroviscous fluid behavior (from Nikolajsen and Hoque, 1990)

practical sense, neither for the ER damper nor for the idea of actively controlled dampers. Therefore there was no real incentive for publication at the time. However, the first author continues to receive requests for information about this research from people who have heard of it. The contemporary enthusiasm for active control of almost every thing is not lost on bearing researchers, especially in the academic community. It begs a description of the challenges faced, the lessons learned, and the disappointments encountered in the study to be reported here, and in the paper describing the experiments [5].

Most previous investigators considered the ER damper to produce a purely Coulomb type of damping force and this was the assumption initially used by the present authors in this study, before the experiments were performed.

Bearing Damper Requirements for Aircraft Engines

Squeeze-film dampers (SFD) have been used extensively in almost all aircraft turbines designed since 1970 to damp imbalance response, and are probably a major contributor to the rarity of rotordynamic instability encountered in these engines. However, the SFD has some shortcomings that are difficult to overcome. For example, recent testing of an aircraft engine rotor-bearing system by the first author showed a large increase in rotor response when the oil temperature was raised from 80°F to 247°F. According to well established SFD theory, the damper actually increases the response whenever the local rotor imbalance exceeds 2.3 times the damper clearance. Also, nonlinear phenomena (bistable jump up) has been reported and extensively investigated [6]. Twenty-five years of applied SFD research has failed to produce reliable design tools that can accurately predict the performance of any except the simplest geometry under laboratory conditions. Realistic conditions include high speeds and low supply pressures relative to the peak pressure in the oil film. Cavitation and air entrainment under these conditions produce a two phase air-oil mixture that greatly reduces the direct damping coefficient [7,8] and that is not highly repeatable even under apparently identical operating conditions.

Rotor damping at the bearings is most needed near critical speeds and is relatively ineffective at other speeds as shown by the classic family of response curves in Fig. 3. The three curves are for three different values of damping. Rotor damping increases the dynamic bearing loads at supercritical speeds as shown in Fig. 4, so an ideal damper should have its damping disappear at speeds where the viscous forces are deleterious. The SFD actually has this characteristic due to the air entrainment, a feature that has not been widely appreciated. To the contrary, several researchers including the first author have suggested in the past that raising the supply pressure could make the SFD more effective and more

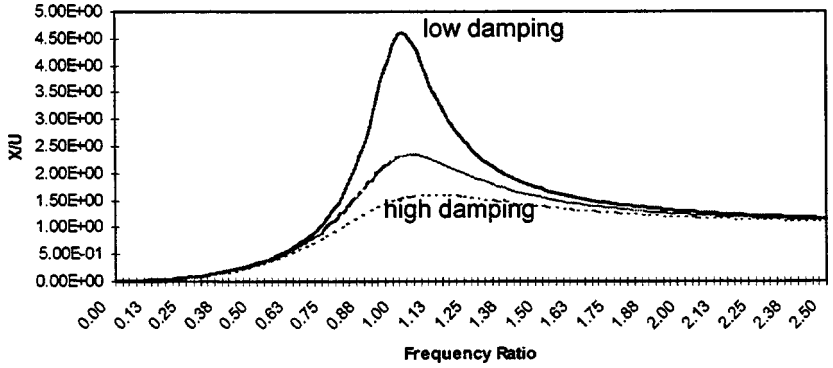


Fig. 3 Effect of viscous damping on vibration amplitude

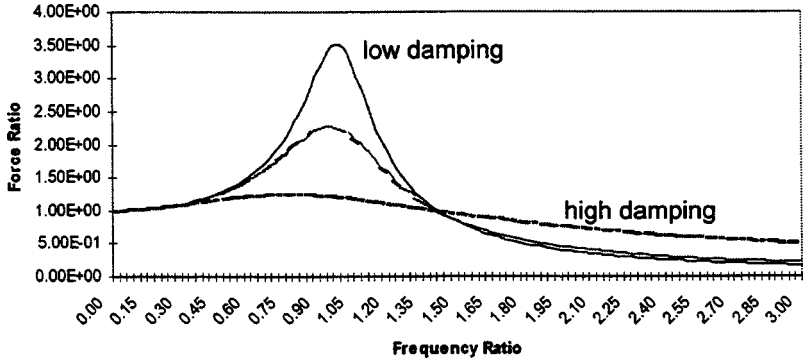


Fig. 4 Dimensionless bearing force vs. frequency ratio

Table 1 Parameters of the XTE-45 engine

Bearing Box Parameter	#1 bearing	#3 bearing
Length	12.7 cm (5")	15.24 cm (6")
Max. Diameter	33.02 cm (13")	38.1 cm (15")
Critical and max. speeds	4700 rpm, 11,500 rpm	11,000 rpm, 13,600 rpm
Operating temperature	121° C (250° F)	177° C (350° F)
Radial displacement	.0356 mm (.0014")	.0305 mm (.0012")
Desired damping	70 kN-s/m (400 lb-s/in)	70 kN-s/m (400 lb-s/in)

predictable by eliminating cavitation [7,8]. But this would lose the desirable characteristic of low bearing forces at high speeds.

Active Dampers for Aircraft Engines

The first experiments with the ER damper by Nikolajsen and Hoque [2] demonstrated that the ER damper has a high damping capability and offers controllability of the rotor response at the first critical speed (which was a rigid body pitching mode in their experiments). One year later a funded research project titled ''Development of Active (Variable Rate) Rotor System Dampers'' was carried out to investigate the feasibility of active controlling bearing damping in aircraft engines. The ER damper concept with the control system was eventually selected as the prototype active damper from seven active damper concepts (SFD, Hydrostatic SFD, ER damper, Piezoelectric Pusher Damper, DC Electromagnetic Damper, AC Electromagnetic Damper and Gas Damper).

The No. 1 and No. 3 bearings of General Electric's XTE45-1 engine were chosen for evaluating the seven candidate devices including the ER damper. The XTE-45 was an experimental aircraft gas turbine engine with dual rotors, high pressure (HP) and low pressure (LP). It was an early test bed for the ''Integrated High Performance Turbine Engine Technology'' (IHPTET) program, a long-range development project jointly funded by government and industry [9]. Bearings 1 and 3 are on the LP rotor and HP rotor respectively.

One of the important criteria in selecting a damper was sizing, to fit within an available envelope and still be able to produce the desired bearing forces. The engine used for the feasibility study had approximate design parameters and requirements for bearing dampers as shown in Table 1. Like most real aircraft engines, the XTE-45 ran through more than one critical speed in its operating speed range. There are some speeds where damping is desirable and other speeds where it is not. The desired damping in Table 1 refers to the value of equivalent viscous damping at the critical speed. Bearing damping (and radial stiffness) in any aircraft engine also becomes more desirable under maneuver and shock loads (e.g., landings). The available restraint force in aircraft maneuvers is also important (see [10]). Thus, the concept of a damper with controllable forces appears attractive.

Control Schemes for Bearing Dampers

Controllable bearing dampers can be conveniently categorized into the following two classes: Class I has indirect control in the sense that a design parameter (such as the viscosity in a passive viscous damper) is varied based on intermittent measurements of rotor vibration (or speed if the response curve is known a priori). For example, the rotor orbit amplitude might be measured once per rotor revolution and the measurement used to control a linear damping coefficient. A more appropriate name for the Class I type might be parametric control. Class II has instantaneous forces on the bearing directly and continuously controlled by feedback signals from the rotor motion with very rapid response to the signals. Both classes were initially included in the study reported here, but the Class II candidates were dropped early on because of (1) challenging requirements for rapid data acquisition and rapid actuator response, (2) stability issues, and (3) the expected higher cost.

The squeeze film damper was a Class I candidate because it had been found in previous research that the direct damping coefficient is controllable with the oil supply pressure to the damper [7,8]. Increasing the supply pressure reduces the air entrainment and cavitation, thus increasing the damping. However, the precision with which this control could be achieved with variable speed was (and still is) questionable and this was considered to be a major disadvantage of the SFD as a controllable device. The hydrostatic SFD would require a high pressure oil pump and had never been tested in the laboratory. It was extensively analyzed [11] but was eliminated from consideration mainly because of a lack of any experimental knowledge about its performance.

The disk type ER damper was an obvious Class I candidate. At the time of the comparison study it was generally believed that it produced Coulomb damping, the strength of which could be easily controlled by varying the voltage across the disks.

If one believes that the ER damper produces Coulomb damping then it is logical to question whether it would be more cost effective to use a damper with dry friction elements, rather than dealing with high electrical voltage and exotic fluid concoctions. Therefore, a dry friction damper with controllable friction force was also considered and a laboratory test rig was built by San Andres (see Fig. 5 and see [4], 1999, for a theoretical analysis). The experiments yielded some good damping results but some of the demonstrated practical disadvantages are (1) wear of the friction surfaces and (2) changes of the friction coefficient with temperature and surface contaminants. Both of these disadvantages would make a reliable control scheme difficult to implement.

As described in the next section, the eddy current damper would be best suited to Class I control. The AC and the DC electromagnetic dampers were Class II candidates since they are essentially magnetic bearings with continuous velocity feedback. The piezoelectric pusher damper was also a Class II candidate and had some stability problems at the time of this study. It was the subject of fairly intensive research by Palazzolo [12]. The gas damper was a Class I candidate. It was controllable by bleedoff air from the compressor and the engine performance penalty from this was unattractive. A prototype was built and demonstrated

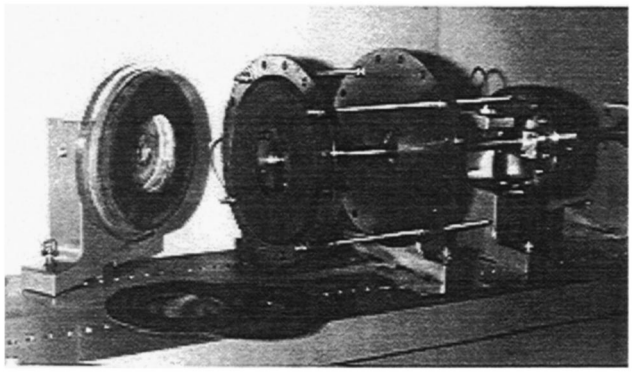


Fig. 5 Coulomb friction test rig with dry friction disks

considerable damping capability in the laboratory [13,14], but the hardware was more than twice as heavy as a traditional SFD for the same size application.

Design Calculations and Test Results for Some of the Dampers

The calculation objective for each of the candidate bearing dampers was to find the lightest and smallest configuration that could meet the conditions of Table 1. Realistic design calculations for DC, AC, and EC-dampers require sophisticated electromagnetic finite element analyses for which the facilities and expertise were not immediately available. Therefore, the field was narrowed through a preliminary evaluation, including a literature survey of the operating principles.

DC-Damper. The actuator of a simple DC electromagnetic damper is shown in Fig. 6. The shaft is supported by a rolling element bearing whose outer race is secured in a ferromagnetic sleeve that is flexibly supported by a squirrel cage and surrounded by electromagnets. Both the sleeve and the electromagnet cores are made of sufficiently thin radial laminations to ensure a flat frequency response up to the maximum rotor speed. The electromagnets are the actuators in a feedback control system (not shown) which senses the shaft vibrations and produces radial forces on the shaft proportional to the velocity of the shaft vibrations, i.e., viscous damping forces. A considerable body of literature exists concerning the design of DC-dampers. Nikolajsen et al. [15] performed one of the first such studies which confirmed very good damping performance, both analytically and experimentally.

The size of DC-dampers, required for the XTE-45, was approximated based on sales literature on DC electromagnetic *bearings* [16]. The resulting damper sizes exceeded the available envelope. The required actuator masses were 16.6 kg and 32.7 kg for the no. 1 and no. 3 bearings, respectively, resulting in specific damping capacities of 4.2 kNs/m/kg and 2.1 kNs/m/kg. The permissible temperature range for the actuator is −250°C to 450°C according to the sales literature. The reliability limited by the control electronics and the power consumption is relatively high due to power electronics and actuator losses. This is a Class II damper.

AC-Damper. This damper looks similar to the DC-damper in Fig. 6 but the ferromagnetic sleeve is replaced with a conducting sleeve (e.g., Aluminum) and the electromagnets are fed with AC power instead of DC power. This results in repulsion between the electromagnets and the sleeve instead of attraction. The advantage is inherent stability, thus, simpler and more reliable electronics and the possibility of Class I operation. The AC-damper is inherently stable because the repulsive force diminishes as it pushes the sleeve away. The DC-damper must be stabilized electronically because the attractive force increases as it pulls the sleeve nearer. Class II operation, however, would be required to ensure adequate damping since the inherent damping capacity of the AC-damper is extremely low. Nikolajsen [17] demonstrated this experimentally for an AC-bearing. Also, the repulsive AC force is only approximately half the size of the attractive DC force. This is due to (1) substantial flux leakage and high reluctance of the magnet-core/sleeve magnetic circuit, (2) the alternating nature of the flux density created by the AC-current, and (3) the resistivity of the conducting sleeve, leading to high eddy-current losses. The required weight and actuator envelope around the XTE-45 bearings were both estimated to be more than twice the DC-damper requirements and the power consumption was estimated to be an order of magnitude higher than for the DC-damper. The AC-damper was rejected on this basis.

EC-Damper. The actuator of a simple EC (eddy current) damper is shown in Fig. 7. The shaft is supported by a rolling element bearing whose outer race carries a conducting disk (e.g., aluminum) which is flexibly supported by a squirrel cage. A ring of DC electromagnets face both sides of the disk. The electromagnets are the actuators in a Class I feedback loop (not shown) which senses the shaft vibrations intermittently and produces a proportional DC-current in the coils which creates a magnetic flux path through the magnet cores and the disk. When the disk vibrates, the magnetic flux lines penetrating it gives rise to eddy-currents in the disk which are directed such that the resulting electromagnetic forces try to prevent the vibrations. The result is radial restoring forces on the disk which are proportional to the velocity of the vibration, i.e., viscous damping forces. A considerable body of literature exists on the design of EC-dampers, see for example Cherry [18]. One obvious advantage of both the EC-damper and the ER-damper over the DC and AC-dampers and SFD for aircraft applications is that the damping capacity does not deteriorate at large radial displacements. This is useful for main-

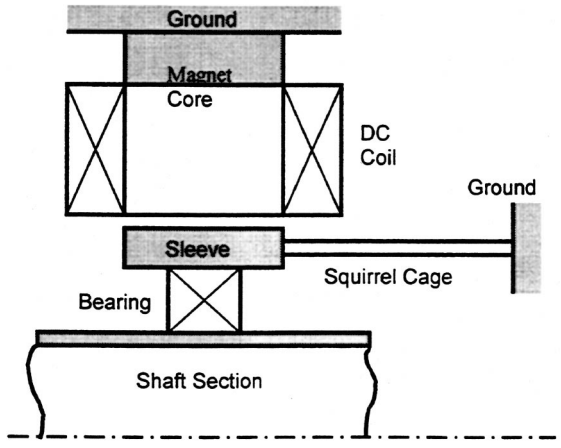


Fig. 6 DC electromagnetic damper operating principle

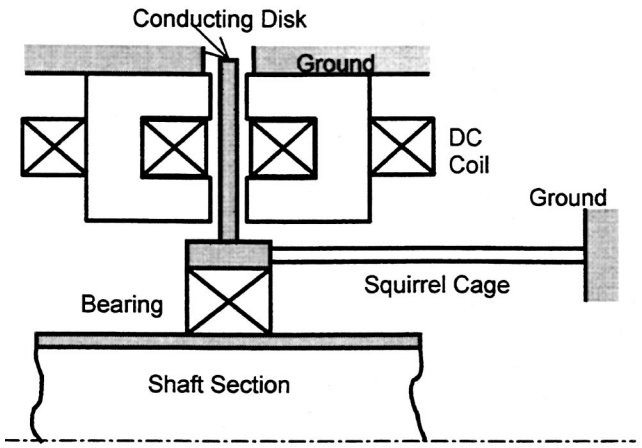


Fig. 7 Eddy-current damper operating principle

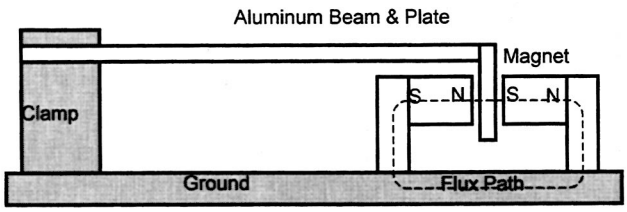


Fig. 8 Eddy-current damper rig

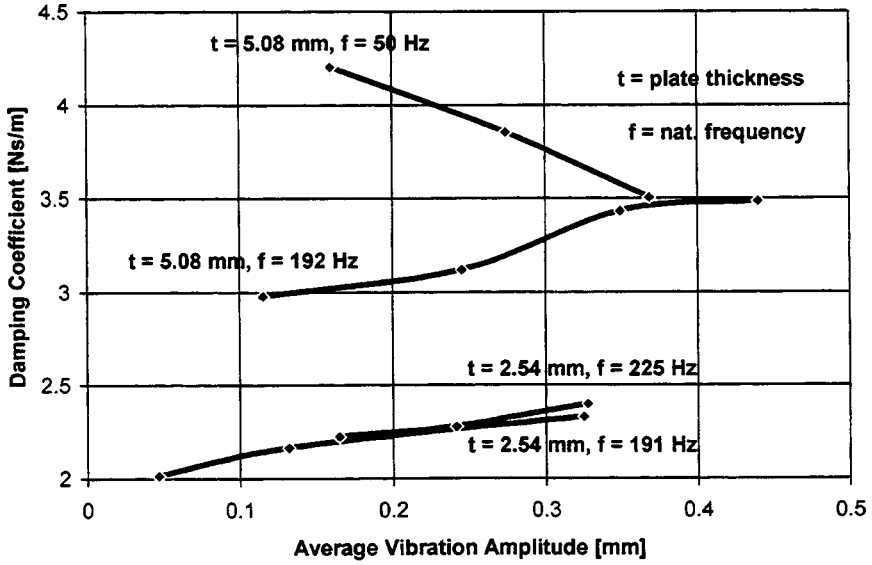


Fig. 9 Measured eddy current damping

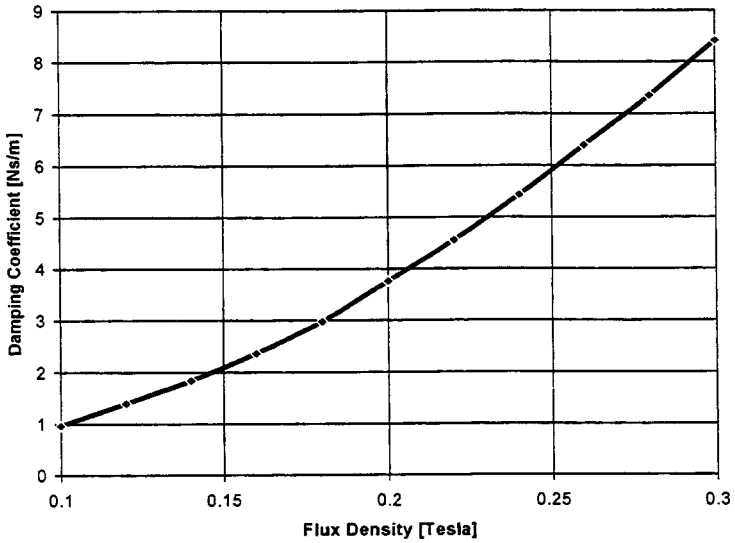


Fig. 10 Predicted eddy current damping

taining the damping capacity during shock loading (e.g., hard landing) and after a blade loss when high and potentially destructive vibration amplitudes need to be attenuated.

The size of the EC-dampers needed for the XTE-45 was approximated based on experimental work with a simple EC-damper rig, see Fig. 8. The cantilevered Aluminum beam has a perpendicular aluminum plate attached at the free end. The plate moves freely in the air gap between the permanent magnets, cutting the magnetic flux lines when the beam vibrates in its first bending mode. Thus, eddy-currents are generated in the plate resulting in viscous drag on the plate. The natural frequency (f_n) in Hz and the logarithmic decrement of the first bending mode are measured by impact testing. The beam is regarded as a 1 D.O.F. system such that the damping coefficient (c) can be written as

$$c = \zeta k/(\pi f_n), \tag{1}$$

where ζ is the critical damping ratio, and

$$k = 3EI/l^3 \tag{2}$$

is the static beam stiffness. The eddy-current damping is determined by subtracting the results from testing with and without the magnets.

Beams of different lengths were used to produce natural frequencies f_n of 50 Hz, 192 Hz, and 225 Hz corresponding to 3000 rpm, 11,500 rpm, and 13,500 rpm. Two plate thicknesses of 2.54 mm and 5.08 mm were used. The vibration amplitude was kept of the order of 0.25 mm to simulate gas turbine engine conditions. The results are shown in Fig. 9. Note the following typical characteristics of EC-damping:

- The damping coefficient increases with vibration amplitude (except at 50 Hz where the accelerometer sensitivity is believed to have deteriorated to the extent where the trend is considered unreliable).
- The damping coefficient increases with plate thickness. The depth of penetration of eddy-currents into the plate at 250 Hz was estimated to be about 6 mm. Therefore, the plate needs to be at least 6 mm thick to allow maximum eddy current generation, thus, maximum damping.

• The damping coefficient decreases with increasing vibration frequency. The damping reduction is about 15 percent from 50 Hz to 192 Hz.

The damping coefficient was also estimated numerically as a function of flux density for the experimental damper design with the 5.08 mm plate thickness, Choi and Nikolajsen [19]. This required a number of simplifying assumptions, including no reduction in damping with increasing frequency. The results are shown in Fig. 10. The flux density in the airgap of the experimental damper was measured to be approximately 0.2 Tesla, thus, according to Fig. 10, the damping coefficient should be 3.7 Ns/m, which is surprisingly close to the measured values.

These results were used to estimate the size of an EC-damper capable of producing a damping coefficient of 70 kNs/m at 15,000 rpm. The best estimate gave an actuator mass of 907 kg, thus a specific damping coefficient of 77 Ns/m/kg, which is unacceptably low, and the damper had to be rejected on this basis alone.

Further background on the electrically related damper principles discussed in this paper can be found in Nikolajsen and Palazzolo [20].

The ER Damper—The Final Choice

Detailed design calculations for the ER damper are lengthy and can be found in Ying [10]. Assuming aluminum construction and an ER fluid shear strength of 8274 Pa (1.2 psi), the specific damping is calculated to be 48.9 kN-s/m per kg (127 1b-s/in per pound) and 26.6 kN-s/m per kg (69 1b-s/in per pound) for the number 1 and 3 bearings, respectively. The total damper weights calculated for each bearing are 1.45 kg (3.2 1b) and 2.63 kg (5.8 1b), respectively. Figures 11 and 12 show the damper weight and the specific damping versus ER fluid shear strength calculated with the assumption of pure Coulomb damping represented by equivalent viscous damping. They lead one to conclude that shear strength is the most important characteristic of an ER fluid for this application.

These results indicated that the ER damper is far superior to any of the Class I candidates except the traditional squeeze film damper with variable supply pressure, and the latter has questionable control capability. The main concerns about the ER damper were of a type that could only be answered by testing, such as high voltage arcing, the true nature of the damping model, and ER fluid robustness. Hence, the ER damper was chosen to explore further with laboratory experiments in a rotordynamic test rig (see [5]).

Conclusions of This Study

1 Controllable bearing dampers for aircraft engines appear to offer some advantages, although the traditional squeeze film damper without control has an intrinsic desirable characteristic that automatically reduces bearing forces at high speeds. The weight, size, and cost penalties for all of the damper design candidates considered are probably unacceptable when compared to the passive squeeze film damper as commonly found today in most aircraft engines.

2 The ER damper invented by Nikolajsen was found to meet the technical requirements of the designated aircraft engine and had already been shown to be controllable by electric voltage. The main concerns about it were high voltage arcing and ER fluid robustness. The ER damper was chosen for further study and experimental evaluation with active control.

3 Assuming the Coulomb model is correct, the most important parameter of an ER fluid for producing required levels of damping

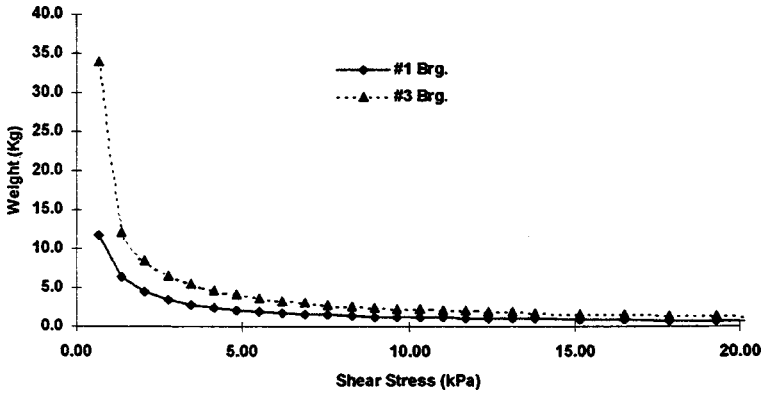


Fig. 11 ER damper weight vs. yield stress of the ER fluid

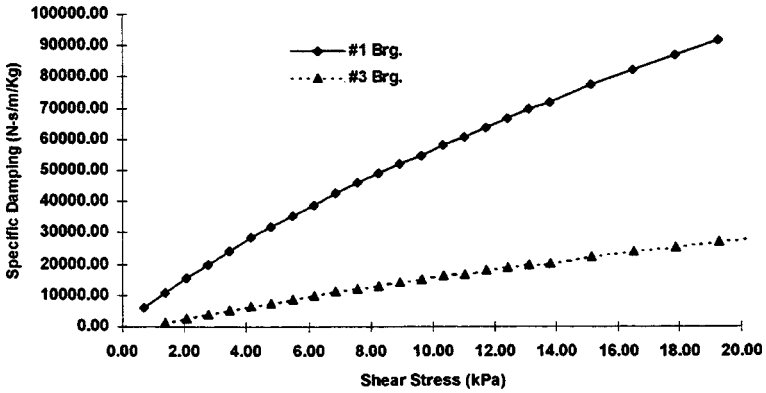


Fig. 12 Specific damping of the ER damper vs. yield shear stress of the ER fluid

is the yield shear stress with voltage applied. But the experiments performed after this study [5] show that the voltage also changes the viscosity of the fluid.

Acknowledgment

These active damper studies were sponsored at Texas A&M University by the General Electric Co. under the capable direction of Mr. Al Storace. His contributions to this work are gratefully acknowledged.

References

[1] Nikolajsen, J. L., and Hoque, M. S., 1988, ''An Electroviscous Damper,'' ''*Rotordynamic Instability Problems in High-Performance Turbomachinery*,'' NASA CP 3026, pp. 133–141.
[2] Nikolajsen, J. L., and Hoque, M. S., 1990, ''An Electroviscous Damper for Rotor Applications,'' ASME J. Vibr. Acoust., **112**, pp. 440–443.
[3] Duff, A. W., 1896, ''The Viscosity of Polarized Dielectrics,'' Phys. Rev., **4**, No. 1, pp. 23–38.
[4] Vance, J. M., and San Andres, L., 1999, ''Analysis of Actively Controlled Coulomb Damping for Rotating Machinery,'' presented at the ASME Turbo-Expo '99 Conference, Indianapolis, IN.
[5] Vance, J. M., and Ying, D., 1999, ''Experimental Measurements of Actively Controlled Bearing Damping With an Electrorheological Fluid,'' presented at the ASME Turbo-Expo '99 Conference, Indianapolis, IN.
[6] Vance, J. M., 1988, *Rotordynamics of Turbomachinery*, Wiley, New York, p. 243.
[7] Zeidan, F. Y., and Vance, J. M., 1989, ''Cavitation Leading to a Two Phase Fluid in a Squeeze Film Damper,'' STLE Tribol. Trans., **32**, pp. 100–104.
[8] Zeidan, F. Y., and Vance, J. M., 1990, ''Cavitation and Air Entrainment Effects on the Response of Squeeze Film Supported Rotors,'' ASME J. Tribol., **112**, pp. 347–353.
[9] Valenti, M., 1995, ''Upgrading Turbine Engine Technology,'' Mech. Eng., **117**, pp. 56–60.
[10] Ying, D. Z., 1993, ''*Experimental Study of an Electrorheological Bearing Damper With a Parametric Control System*,'' Master of Science thesis, Mechanical Engineering, Texas A&M University, College Station, TX.
[11] San Andres, L. A., 1992, ''Analysis of Hydrostatic Journal Bearings With End Seals,'' ASME J. Tribol., **114**, pp. 802–811.
[12] Palazzolo, A. B., et al., 1989, ''Piezoelectric Pushers for Active Vibration Control of Rotating Machinery,'' ASME J. Vibr. Acoust., **111**, pp. 298–305.
[13] Sundararajan, P., and Vance, J. M., 1995, ''A Theoretical and Experimental Investigation of a Gas-Operated Bearing Damper for Turbomachinery: Part I—Theoretical Model and Predictions,'' ASME J. Eng. Gas Turbines Power, **117**, No. 4, October, pp. 742–749.
[14] Sundararajan, P., and Vance, J. M., 1995, ''A Gas-Operated Bearing Damper for Turbomachinery—Theoretical Predictions versus Experimental Measurements: Part II—Experimental Results and Comparison With Theory,'' ASME J. Eng. Gas Turbines Power, **117**, No. 4, pp. 750–756.
[15] Nikolajsen, J. L., Holmes, R., and Gondhalekar, V., 1979, ''Investigation of an Electromagnetic Damper for Vibration Control of a Transmission Shaft,'' Proc. MechE, **193**, pp. 331–336.
[16] *Active Magnetic Bearings*, 1986, Magnetic Bearings Inc., Radford, VA.
[17] Nikolajsen, J. L., 1989, ''Experimental Investigation of an Eddy-Current Bearing,'' in *Magnetic Bearings*, G. Schweitzer, ed., Springer, New York, pp. 111–118.
[18] Cherry, L. B., 1960, ''Electro-Magnetic Induction Damping of Vibratory Motion,'' Noise Control, **Sept/Oct**, pp. 8–11.
[19] Choi, K. N., and Nikolajsen, J. L., 1989, ''The Effect of Vibration Frequency in Eddy Current Dampers,'' internal report, Texas A&M University, College Station, TX.
[20] Nikolajsen, J. L., and Palazzolo, A. B., 1992, ''Electromagnetic Force Application in Rotordynamic Research,'' in Tani, J., and Takagi, T., eds., *Electromagnetic Forces and Applications*, Elsevier, New York, pp. 163–166.

Dynamic Coefficients of Stepped Labyrinth Gas Seals

Klaus Kwanka
Chair of Thermal Power Systems,
Technische Universität München,
Munich, Germany

Rotor-fluid interactions can cause self-excited shaft vibrations of high density turbomachinery. Often the amplitude of the vibrations reaches unacceptably high amplitudes and the scheduled power or running speed cannot be achieved. One of the most important sources of excitation is the flow through labyrinth seals. For a reliable design it is necessary to predict these forces exactly, including not only stiffness but also damping coefficients. As the forces in labyrinth gas seals are rather small only minimal experimental data is available for the comparison and validation of calculations. Meanwhile a new and easy-to-handle identification procedure enables the investigation of numerous seal geometrys. The paper presents dynamic coefficients obtained with a stepped labyrinth and the comparison with other seal concepts. [S0742-4795(00)00903-0]

Introduction

High-density turbomachines often operate close to the stability limit which is characterized by a violent vibration behavior. In case of instability, the vibration due to unbalance is overlaid by a rapidly increasing vibration close to an eigenfrequency of the rotor. The unstable vibrations are of the self-excited type and they are mainly generated by hydrodynamic bearings and by the flow of the working fluid in the turbomachine. All exciting mechanisms are induced by lateral forces which transfer rotative energy into the rotor bending vibrations.

A mechanism which leads to fluid-generated forces, so-called steam whirl, was first described by Thomas [1]. The forces are traced to non-uniform leakage losses which cause variable blade forces and result in a lateral and therefore exciting force. Alford [2] discovered an additional effect which may increase the excitation. When the blades are shrouded, an asymmetric pressure distribution in relation to the rotor deflection results in a second excitation component. Meanwhile, it turned out that one of the main parameters influencing the asymmetry of the pressure distribution is the swirl of the flow at the entrance to the labyrinth seal.

The rotor-fluid interaction in gas seals is usually described by a linearized relation. The inertia coefficients neglected here are considered only in case of liquid seals (Childs [3]).

$$\overline{F} = -\begin{pmatrix} K & k \\ -k & K \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} C & c \\ -c & C \end{pmatrix}\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} \tag{1}$$

With small motions around the centered position it can be assumed that the matrices of the stiffness coefficients K and k and the damping coefficients C and c are symmetric in the main diagonal and skew-symmetric in the secondary diagonal. Experimental investigations (Wagner and Steff [4]) indicate, that the simplification mentioned above is justified even for small deviations out of the centered position. The linearity of the equation is given up to an eccentricity ratio of 0.5.

The destabilizing force caused by the cross-coupled stiffness is counteracted by the direct damping; therefore, it is also important to know and to take into consideration the damping. The experimental identification of all the dynamic coefficients requires a motion of the rotating rotor relative to the stator. Most of the test rigs use heavy and stiff rotors and therefore the realization of a predefined motion is an ambitious task. The seal forces influence the dynamic behavior of the system. The response caused by the motion is an unsteady reaction force or the corresponding circumferencial pressure distribution, which must be measured with regards to amplitude and phase [5–9].

Recent arrangements use active magnetic bearings to carry the rotor and to measure the forces simultaneously [10,11]. The latter succeeded in measuring dynamic coefficients even under extreme conditions in pressure and rotational speed [4]. Presumably, only few results of this test facility will be published in the near future and will therefore be accessible to a broader public.

However, only few experimental dynamic coefficients for gas seals are available at the moment. One of the reasons might be that the forces generated in gas seals are at least one order in magnitude smaller than in liquid seals. A survey of the experimental approaches and the investigated seal geometries is given by Childs [3].

Experimental Procedure

The new approach developed by the author tries to avoid some of the restrictions and difficulties caused by the rigid rotor concept [12]. The starting point of the identification procedure is a Jeffcott-type rotor carried in fluid film bearings with a test seal and a magnetic bearing placed at the midspan position (Fig. 1). The shaft is flexible.

The flow through the test seal causes a change in the dynamic behavior of the rotor. The (unknown) nonconservative acting coefficients as the cross-coupled stiffness and the direct damping alter the stability limit of the system. The conservative working coefficients, namely the direct stiffness and the cross-coupled damping, influence the vibrational frequencies. Both, the change in the stability and the frequency can be measured, which finally helps to find the dynamic coefficients.

An important role in the identification procedure is played by the magnetic bearing. It helps to simulate flow-like, displacement-dependent forces, thus enabling an excitation of the rotor up to the stability limit; also, the frequency alteration caused by the flow can be compensated. Note, that the magnetic bearing is not used as a bearing in the traditional sense, but as an excitation source. The actual use of the magnetic bearing was inspired by Ulbrich [13].

Two measurements are necessary to find the changes caused by the flow through the test seal. First, a measurement without flow through and second, a measurement with flow. The first measurement helps to identify the stability behavior of the rotor which is controlled mainly by the dynamic characteristics of the involved hydrodynamic bearings.

$$q_0 = \frac{(C\Omega a. - ka)_{\text{BEARING}}}{a} \tag{2}$$

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-20. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

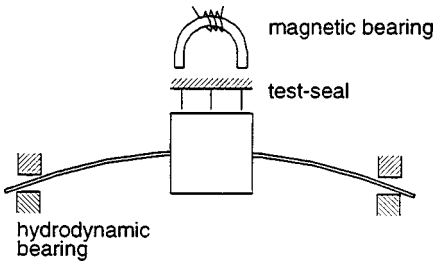


Fig. 1 Jeffcott rotor with test seal and magnetic bearing

The flow through the test-seal alters the stability limit (and the whirl frequency) of the rotor. Again the acting magnetic excitation at the stability limit q can be measured.

$$q = q_0 - q_{\mathrm{SEAL}} \tag{3}$$

The changes in magnetic excitation at the stability limit Δq and the magnetic direct stiffness Δr which is necessary to compensate the change in whirl frequency are known due to a previous calibration of the magnetic bearing. The following relations between the dynamic coefficients in Eq. 1 and the measured changes Δq and Δr exist, when a circular whirling orbit is assumed.

$$\Delta q = q_{\mathrm{SEAL}} = k - C\Omega \tag{4}$$

$$\Delta r = -K - c\Omega \tag{5}$$

The influence of the flow on the stability limit depending on the whirling frequency is displayed in Fig. 2. As the stiffness of the shaft and the whirling frequency increases, less energy is transferred to the vibrational bending mode. The magnetic excitation at the stability limit q of the forward and backward vibration increases, too (continues line). The assumed forward excitation of the flow reduces the stability of the forward mode and elevates the stability of the backward mode (broken line). If the aerodynamic excitation is high, a magnetic stabilization becomes necessary (positive whirl frequency, area below the frequency axis).

When the measurement is repeated for at least two whirl frequencies, then all unknown coefficients can be calculated (Eqs. 4 and 5). The measured changes in magnetic stiffness come to lie on a straight line. The intersection with the ordinate gives the stiffness value and the slope of the line determines the damping value (Fig. 3).

To enhance the fitting reliability of the line needed for identification, in this paper the measurements were repeated for three forward and three backward whirling frequencies. An exact determination of the stability limit is essential in order to be able to obtain accurate coefficients. Usually, this limit can be fixed within ± 1–2 N/mm [14].

The stability limit is characterized by a sudden and distinct change of the vibration amplitude and the orbit shape. In the frequency domain the vibrational signal contains in addition to the rotating frequency the unstable eigen-frequency. The spiral orbit of this band-filtered eigen-vibrations registered in about 0.5 s is documented in Fig. 4.

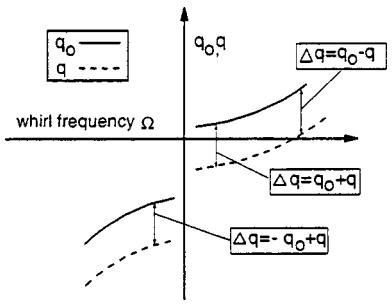


Fig. 2 Change in stability of the forward and the backward mode caused by the seal

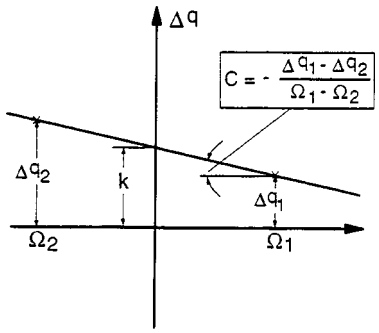


Fig. 3 Identification of the cross-coupled stiffness and the direct damping

The adjusted gain of the control system at the stability limit corresponds to a known cross-coupling (or direct stiffness) of the magnetic bearing. The linearized magnetic characteristics were extracted out of a static and a dynamic calibration procedure. The used configuration for the calibration was identical to that used in the subsequent experimental investigations. The maximum deviation of the force-current factor from the linear least square fit is about 5 percent. Since the vibrational orbit is frequency dependent a variable part of the calibrating range is involved and the average error will be considerably lower than the maximum value mentioned above. One identification run includes measurements with different frequencies and therefore the error will vary within the identification process.

The measurement accuracy and reproducibility were discussed in Kwanka [15]. First measurements with the identification procedure are displayed in Fig. 5 and should give an impression of the fluctuations about the fitted line which are representative for all of the following results [12]. Each of the single dots represent the average of five Δqs measured with one whirl frequency. The used seal geometry is of the staggered type with two cavities. As expected the cross-coupled-stiffness (intersection with ordinate) and the direct damping (slope of the fitted line) increase along with the pressure ratio Δp.

Without difficulties the procedure can be extended for the additional identification of the inertia coefficients. In this case, measurements with at least three whirling frequencies are necessary and a quadratic curve must be fitted.

As expected, the test rig build up to identify the coefficients is completely different from the other arrangements employed in this matter (Fig. 6). The test seals including the casing are located there, at midspan position. Two identical seals are placed sym-

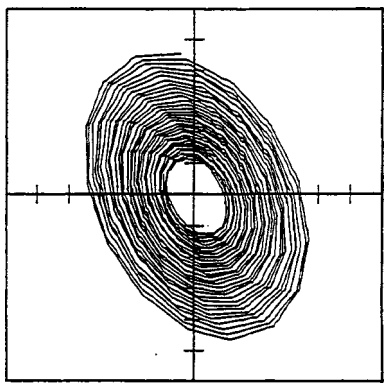


Fig. 4 Spiral orbit of the unstable eigen-vibration

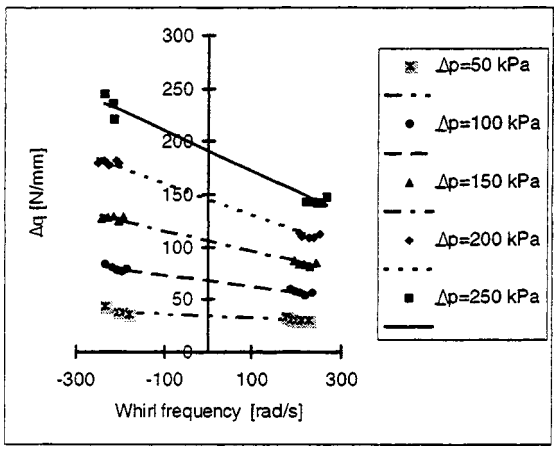


Fig. 5 Identification of the cross-coupled stiffness and the direct damping (dots=measurement; lines=curve fitting)

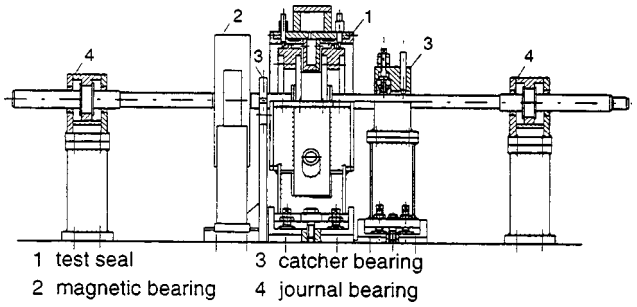


Fig. 6 Test rig for the identification of dynamic coefficients

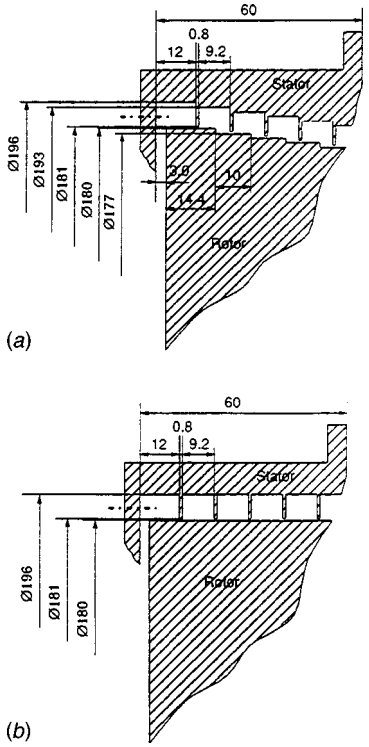


Fig. 7 (*a*) Geometrical data of the stepped labyrinth seal (STS14); (*b*) geometrical data of the tooth on stator labyrinth seal (TOS24).

metrically to the guide vane ring which generates the entry swirl. Both, the seal rotor/stator and the guide vane ring can easily be replaced and changed to another geometry. The arrangement with two identical seals helps to double the small forces and average them at the same time. The magnetic bearing is mounted as close as possible to one site of the seal casing. Because the position of the magnetic bearing differs from the seal-position it is necessary to transfer the forces to a common point of application [16]. The coefficients in Fig. 5 include two seals without transfer.

The rotor with a diameter of 23 mm is carried in fluid film bearings and flexibly connected to a variable d.c. motor. The variation of the frequency is realized by varying the spacing between the bearings. To avoid violent vibrations with high amplitudes in the unstable vibration regime, retainer bearings are placed at both sides of the test-seal housing. After each change in the whirl frequency, which means a dismounting and remounting of the journal bearing pedestals, both measurements (with and without flow through the test-seal) are repeated.

Experimental Results

As mentioned above, the test-seal internals and the guide vane ring can be easily removed and changed to other labyrinth geometries or guide vane angles. In this paper the main emphasis lies on a stepped labyrinth seal (STS14) with four cavities and strips on stator s (Fig. 7(*a*)). The experimental findings are compared to a tooth on stator labyrinth seal (TOS24; Fig. 7(*b*)). Both seals have the same geometrical data in terms of clearance (0.5 mm) and number of cavities (4). The inflow region and the cavity length are identical. The only difference between the seals consists in the step of 1.5 mm, which leads to a reduction of the clearance area from the entrance of the seal to the exit.

The main task of labyrinth seals is still to minimize the leakage flow. First steps to improve the sealing behavior are: tightening the clearance, increasing the number of cavities or hindering in some way the flow in the cavity (e.g., by changing from the look-through to the stepped design). The expected decrease of the leakage of the stepped seal as compared to the looked-through seal is displayed in Fig. 8.

Unfortunately, the leakage flow of the TOS24 seal was measured only with two lower pressure differences. The mass flow through the stepped seal is approximated by a polynomial fit. The

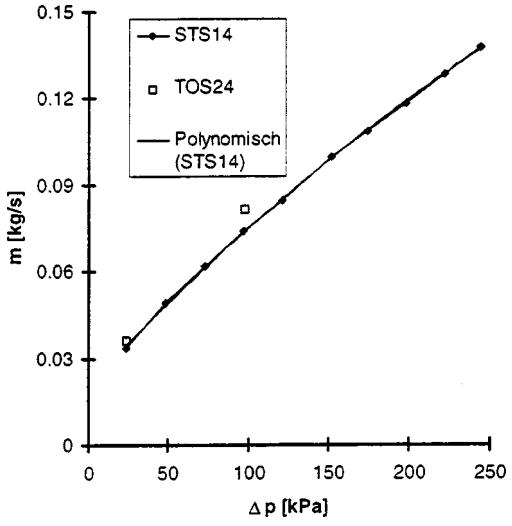


Fig. 8 Leakage loss of labyrinth seal versus the pressure difference (*n*=750 rpm)

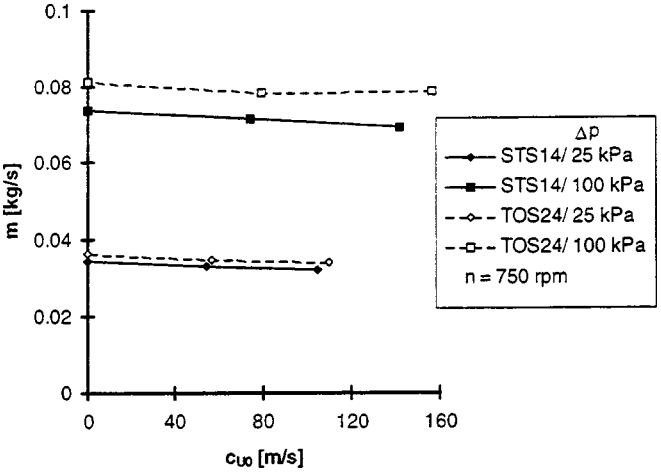


Fig. 9 Leakage rate in dependence of the entry swirl

test medium is compressed air which is cooled, dried and metered by an orifice before entering the seals. The air is expanded to ambient conditions.

The dynamic coefficients are measured in dependence of the entry swirl. When the pressure difference on the seal is held constant, the mass flow through the seal slightly decreases with higher swirl values (Fig. 9). Again, the better sealing performance of the stepped seal becomes obvious.

The nonconservative coefficients have a considerable impact on the dynamics of the rotor. The cross-coupled stiffness excites the rotor, whereas the direct damping counteracts this excitation. The nonconservative coefficients of the stepped seal with two pressure differences are displayed in Fig. 10.

The cross-coupled stiffness shows the usual almost linear behavior in dependence of the swirl condition at the entrance of seal. A higher pressure difference also leads to higher stiffness values. The dependencies of the direct damping on the pressure difference and the swirl is not that clear. Noteworthy is, however, the fact that the flow generates direct damping even without swirl.

The conservative acting coefficients do not influence the stability of the rotor directly, but indirectly via the deflection of the vibrational mode. When the modal deflection increases due to a weakening influence of the seal, then more energy is transferred to the rotor and the stability limit decreases. The stepped labyrinth seal generates a positive direct stiffness, and, therefore, a stabilizing effect (Fig. 11).

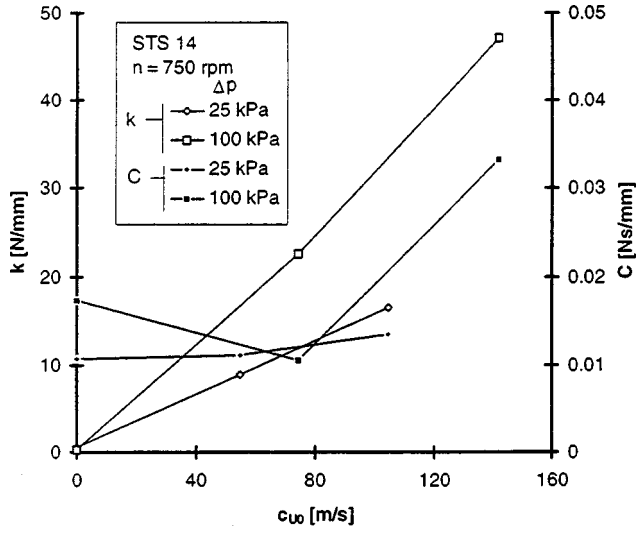


Fig. 10 Nonconservative dynamic coefficients of a stepped labyrinth seal

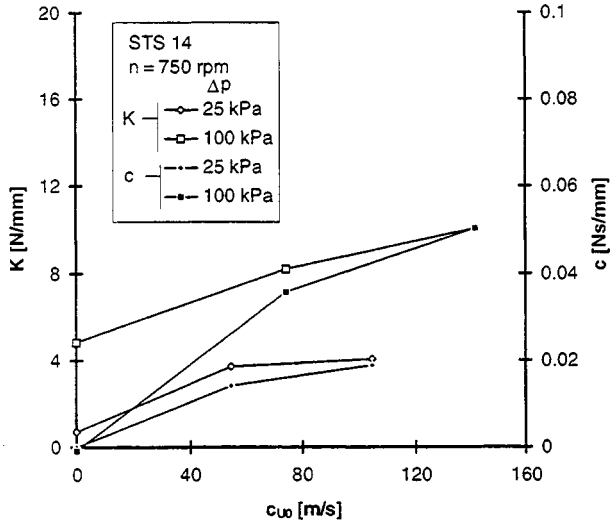


Fig. 11 Conservative dynamic coefficients of a stepped labyrinth seal

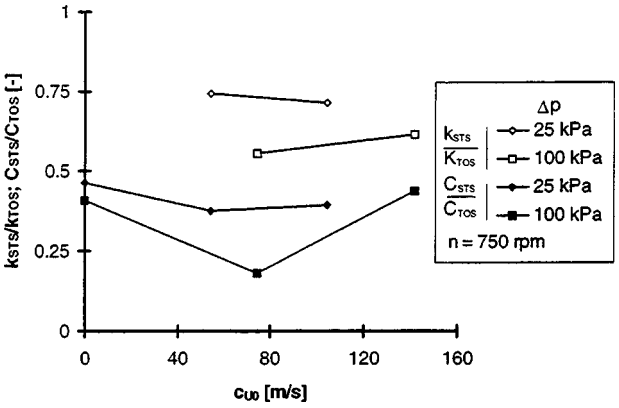


Fig. 12 Comparison of a stepped and a tooth-on-stator seal

Solely the influence on the frequency might be of interest to avoid a coincidence with the operating frequency. An elevation of the pressure difference or the entry swirl on the conservative coefficients leads to a moderate increase.

Finally, a comparison of the nonconservative coefficients of the stepped seal with the very common look-through seal should indicate which seal concept is more suited from the dynamic point of view. In Fig. 12, the cross-coupled stiffness and the direct damping of the stepped seal are related to the numbers of the tooth-on-stator seal. The cross-coupled stiffness of the stepped seal is distinctly smaller; unfortunately, at the same time the reduction of direct damping is more pronounced. As a consequence, the better sealing labyrinth posses the poorer dynamic suitability. This behavior is in some respects comparable with the interlocking seal which also shows low direct damping values [17].

Conclusions

A new easy-to-handle approach for the identification of the dynamic coefficients of labyrinth gas seals allows the direct and systematic comparison of seal concepts. The objective is to obtain an optimized seal concept which combines a good sealing and good a rotordynamic performance.

As expected, the stepped labyrinth seal leaks less than a look-through seal. The dynamic properties of the stepped seal are not that favorable. When stepped labyrinth seals are used it must be ensured that there is still enough damping in the system.

Nomenclature

a = amplitude of the vibrational mode
C,E,K = direct damping, inertia, stiffness coefficients
c,e,k = cross damping, inertia, stiffness coefficients
c_{u0} = entry swirl
F_T = lateral force
Δr = change in direct stiffness of the magnetic bearing
Δp = pressure difference
Δq = change in magnetic excitation at the stability limit
x,y = coordinates of displacement
Ω = whirling frequency
STS14 = nomenclature of the stepped labyrinth
TOS24 = nomenclature of the look-through labyrinth (tooth on stator)

References

[1] Thomas, H.-J., 1958, ''Instabile Eigenschwingungen von Turbinenläufern, angefacht durch die Spaltströmungen in den Stopfbuchsen und Beschaufelungen,'' Bull. de AIM 71, pp. 1039–1063.

[2] Alford, J. S., 1965, ''Protecting Turbomachinery From Self-Excited Rotor Whirl,'' ASME J. Eng. Gas Turbines Power, **87**, pp. 333–344.

[3] Childs, D. W., 1993, *Turbomachinery Rotordynamics*, Wiley, New York.

[4] Wagner, N. G., and Steff, K., 1996, ''Dynamic Labyrinth Coefficients From a High-Pressure Full-Scale Test Rig Using Magnetic Bearings,'' NASA Conference Publication to be published, proceedings of a workshop held at Texas A&M University, College Station, TX, May 6–8, 1996.

[5] Childs, D. W., Nelson, C. E., Nicks, C., Scharrer, J., Elrod, D., and Hale, K., 1986, ''Theory Versus Experiment for the Rotordynamic Coefficients of Annular Gas Seals: Part 1–Test Facility and Apparatus,'' ASME J. Tribol., **108**, pp. 426–432.

[6] Hawkins, L., Childs, D., and Hale, K., 1989, ''Experimental Results for Labyrinth Gas Seals with Honeycomb Stators: Comparison to Smooth-Stator Seals an Theoretical Predictions,'' ASME J. Tribol., **111**, pp. 161–168.

[7] Yu, Z., and Childs, D., 1996, ''A Comparison of Rotordynamic Coefficients and Leakage Characteristics for Hole-Pattern Gas Damper Seals and Honeycomb Seal,'' NASA Conference Publication to be published, proceedings of a workshop held at Texas A&M University, College Station, TX, May 6–8, 1996.

[8] Wright, D. V., 1983, ''Labyrinth Seal Forces on a Whirling Rotor, ASME Applied Mechanics Division,'' Proceedings of a Symposium on Rotor Dynamical Instability, Vol. 55, Adams, M. L., Jr., ed., ASME, New York, pp. 19–31.

[9] Millsaps, K. T., and Martinez-Sanches, M., 1993, ''Rotordynamic Forces in Labyrinth Seals: Theory and Experiment,'' NASA Conference Publication 3239, proceedings of a workshop held at Texas A&M University, College Station, TX, pp. 179–207.

[10] Matros, M., Neumer, T., and Nordmann, R., 1994, ''Identification of Rotordynamic Coefficients of Centrifugal Pump Components Using Magnetic Bearings,'' Preprints of the Fifth International Symposium on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-5), May 8–11, 1994, Kaanapali, Hawaii, USA, pp. 55–72.

[11] Wagner, N. G., and Pietruszka, W. D., 1988, ''Identification of Rotordynamic Parameters on a Test Stand With Magnetic Bearings,'' Magnetic Bearings Proceedings of the First International Symposium, ETH Zürich, Schweitzer, G., ed., Springer Verlag, New York, pp. 289–299.

[12] Kwanka, K., and Mair, R., 1995, ''Identification of Gas Seal Dynamic Coefficients Based on the Stability Behavior of a Rotor,'' *Proceedings of the 1st European Conference of Turbomachinery—Fluid Dynamic and Thermodynamic Aspects*, March 1–3, Erlangen-Nürnberg, Germany, VDI-Report 1186, pp. 297–309.

[13] Ulbrich, H., 1988, ''New Test Techniques Using Magnetic Bearings,'' Magnetic Bearings Proceedings of the First International Symposium, ETH Zürich, Schweitzer, G., ed., Springer Verlag, New York, pp. 281–288.

[14] Kwanka, K., and Nagel, M., 1996, ''Experimental Rotordynamic Coefficients of Short Labyrinth Gas Seals,'' NASA Conference Publication to be published, proceedings of a workshop held at Texas A&M University, College Station, TX, May 6–8, 1996, pp. 135–144.

[15] Kwanka, K., 1997, ''Rotordynamic Impact of Swirl Brakes on Labyrinth Seals with Smooth or Honeycomb Stators,'' ASME Paper 97-GT-232.

[16] Kwanka, K., 1995, ''Variation of Fluid Flow Forces in Seals With Rotor Bending,'' DE-Vol. 84-2, 1995 Design Eng. Technical Conf., Vol. 3-Part B, ASME, New York, pp. 1277–1282.

[17] Kwanka, K., 1998, ''Influence of Labyrinth Seal Geometry on Rotordynamic Coefficients,'' *Proceedings of the Fifth International Conference on Rotor Dynamics (IFToMM)*, Irretier, H., and Nordmann, R., ed., Vieweg Verlag, pp. 241–251.

Feasibility of Contact Elimination of a Mechanical Face Seal Through Clearance Adjustment

Min Zou

Joshua Dayan

Itzhak Green

The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0405

The feasibility of eliminating contact in a noncontacting flexibly mounted rotor (FMR) mechanical face seal is studied. The approach for contact elimination is based on a parametric study using FMR seal dynamics. Through clearance adjustment it is possible to reduce the maximum normalized relative misalignment between seal faces and, therefore, eliminate seal face contact. Clearance is measured by proximity probes and varied through a pneumatic adjustment mechanism. Contact is determined phenomenologically from pattern recognition of probe signals and their power spectrum densities as well as angular misalignment orbit plots, all calculated and displayed in real-time. The contact elimination strategy is experimentally investigated for various values of stator misalignment and initial rotor misalignment. Contrary to intuition but compliant with the parametric study, the experimental results show that for the seal under consideration contact can be eliminated through clearance reduction. [S0742-4795(00)01503-9]

Introduction

Mechanical face seals are usually used for gas or liquid sealing of rotating shafts. Typical applications can be found in gas turbines, pumps, and compressors. Two types have evolved in mechanical seal design: contacting and noncontacting mechanical face seals. Contacting seals are designed to operate with face contact to minimize leakage at a cost of relatively large friction and wear of the seal faces. Noncontacting seals are designed to operate with certain face separation to reduce frictional heat generation and wear at a cost of some leakage. Statistics of industrial mechanical seals shows that most of them fail prematurely rather than wear out in a normal way [1]. Seal failure comprises a large portion of plant maintenance costs [2]. The state of the art has not proven adequate in guaranteeing noncontacting operation of the most carefully developed mechanical seals, not even for well-defined applications, such as pump seals in the nuclear power industry. Various explanations have been offered to account for unpredictable and premature seal failure, but some basic factors not yet fully understood appear to be the cause. A problem likely to cause seal failure is intermittent contact between the faces. Therefore, contact elimination is of prime importance in preventing seal premature failure, especially in critical applications where seal failure may have severe implications.

Experimental investigation of the dynamics of mechanical seals dates back two decades. Etsion and Burton [3] tested a seal under eccentric loading (i.e., initial stator misalignment) that resulted in ''self-excited oscillations.'' Metcalfe [4] observed seal dynamic whirl at stability threshold to be close to half of the shaft speed in a well-aligned mechanical seal. Sehnal et al. [5] investigated the effects of face coning on the seal performance by comparing torque, face temperature, leakage, and wear of a conventional flat-face seal and of coned face seals. Etsion and Constantinescu [6] found that the stator misalignment and its phase shift are time dependent. Eddy current proximity probes have been the preferred means of measuring the stator motion. All these investigations have been, however, restricted to noncontacting mechanical seals having a flexibly-mounted stator (FMS), where data has been analyzed postmortem in the time domain.

The flexibly mounted rotor (FMR) seal is as prevalent in the industry as the FMS seal. In fact, theoretical work by Green [7,8] has proven that the FMR seal has a superior dynamic behavior compared to the FMS seal. The only experimental investigation available on the dynamic behavior of a FMR seal is that by Lee and Green [9–11]. Results obtained from their test rig showed that the FMR seal was vulnerable to higher harmonic oscillations having superimposed signals of integer multiples of the shaft rotating frequency. Periodic face contact has been determined to be the prime source of such oscillations leading eventually to face wear and imminent failure. After elimination of the higher harmonic oscillations through seal redesign it was found that the deviations between analytical and experimental results did not exceed 15 percent. But, as with all previous experimental investigations, data had been acquired first and analyzed off-line later.

For the purpose of monitoring the seal dynamic behavior in real-time Zou and Green [12] have incorporated a data acquisition system into a personal computer that is capable of acquiring and processing data in real-time by a dedicated on-board processor. They have also introduced the concept of orbit plots for the angular response capable of distinguishing between noncontacting and contacting operation modes. Later, Zou and Green [13] applied a PI control strategy using a personal computer and the same board, not only for data acquisition, but also for control. They have been able to maintain and/or adjust the seal clearance at any desired value, whether fixed or varying. However, the elimination of contact once detected, using clearance adjustment has never been attempted. The purpose of this work is to investigate first analytically and then experimentally whether such a concept is feasible.

Several other studies have dealt with the development of controlled mechanical face seal systems [14–17]. They all concentrated on clearance adjustment through temperature control. Temperatures measured by thermocouples were used as feedback, assuming that these temperatures are directly related to clearance. In reality, the thermocouples measure local temperatures only close to the sealing dam. Adding the likelihood of large thermal inertia, this approach may be vulnerable to large time delays between event occurrence and control action. Obviously, the temperature is also not a direct measure of the clearance; therefore, it may not be the most effective feedback signal for adjusting it.

Seal face contact is usually caused by large relative misalign-

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-147. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

ment between the seal rotor and stator. The study of seal dynamics can help in understanding face noncontacting operation and face separation, and it may suggest a novel strategy of contact control. The present work provides important theoretical insight into the rotordynamics of the noncontacting FMR mechanical face seal utilizing closed-form solutions to assist the strategy of contact elimination. First, the parametric analysis by Zou et al. [18] is applied to the present seal, which affirmatively concludes that contact can be eliminated through variation of the seal clearance. Finally, the feasibility of the proposed contact elimination is investigated experimentally.

Parametric Analysis

Seal Angular Response. The test rig subject for this study is shown in Fig. 1 and is discussed in the next section. The analysis and notation that pertain to the current work are consistent with both the theoretical work of Green [7,8] and the experimental work utilizing the test rig. For conciseness this is not repeated. Only the final results necessary for the current study are summarized below.

In FMR seals the stator misalignment, γ_s, and the initial rotor misalignment, γ_{ri}, always exist due to assembly and manufacturing tolerances, or will eventually develop in time because of shaft deflection and machine deterioration. Both misalignments act as forcing functions to the seal angular modes. The maximum rotor steady-state response to these forcing functions is obtained in terms of static transmissibility, γ_o/γ_s, and dynamic transmissibility, γ_{rI}/γ_{ri}, as follows [7]

$$\frac{\gamma_o}{\gamma_s}=\sqrt{\frac{K_s^2+(D_s\omega)^2}{(K_f+K_s)^2+\left(D_s+\frac{1}{2}D_f\right)^2\omega^2}} \tag{1}$$

$$\frac{\gamma_{rI}}{\gamma_{ri}}=\sqrt{\frac{K_s^2}{[(I_z-I_t)\omega^2+(K_f+K_s)]^2+\left(\frac{1}{2}D_f\omega\right)^2}}, \tag{2}$$

where γ_o is the relative misalignment caused by γ_s only, and γ_{rI} is the relative misalignment caused by γ_{ri} only. K_f and D_f are the fluid film angular stiffness and damping coefficients, respectively [19]. K_s and D_s are the rotor support angular stiffness and damping coefficients, respectively [10]. I_z and I_t are the rotor polar and transverse moment of inertia, and ω is the shaft speed.

Assuming that superposition is valid for small perturbations about the working point (the nominal design point at steady-state), the maximum total relative misalignment between the rotor and stator, $\gamma_{\max}$, is the sum of the rotor responses to both the stator misalignment and the initial rotor misalignment:

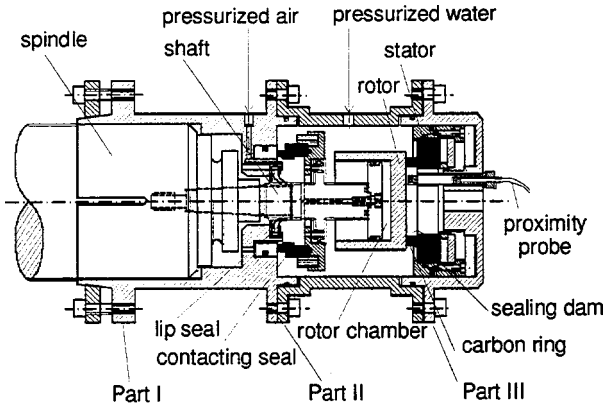


Fig. 1 Schematic of the FMR noncontacting mechanical seal assembly

$$\gamma_{\max}=\gamma_o+\gamma_{rI}. \tag{3}$$

When this relative misalignment is too large it will cause face contact. A basic requirement to avoid contact from the onset is to ensure hydrostatic stability via positive angular film stiffness [19]. In which case, a properly designed seal must have a normalized coning angle, $\beta^*=\beta r_o/C_o$, that is greater than the critical value, $\beta^*_{\text{critical}}=1/R_i$ (where $R_i=r_i/r_o$ is the dimensionless inner radius). Should face contact still occur because of a large dynamic angular response, it will take place at the minimum film thickness at the inner radius. Mathematically, contact is determined whenever the maximum relative misalignment between the seal faces, $\gamma_{\max}$, exceeds a critical value. In a normalized form this condition is [7]:

$$\gamma^*_{\max}=\frac{\gamma r_o}{C_o}>\gamma^*_{\text{critical}}=\frac{1}{R_i}, \tag{4}$$

where r_o is the seal outside radius, and C_o is the clearance. Thus, seal design, manufacturing and operation should ensure $\gamma^*_{\max}<\gamma^*_{\text{critical}}$ at all times in order to avoid face contact.

Parameter Effects. In most applications, the operating conditions of the mechanical seal are not precisely known. Disturbances and variations of operating conditions and the lack of adjustment mechanism to counteract them may cause face contact, ultimately leading to face wear and seal failure. Thus, the incorporation of an active control mechanism may reduce the danger of failure and increase seal reliability. Studying the effects of different input variables on the total normalized relative misalignment may suggest possible control parameters for a contact elimination strategy.

The parameters of Eqs. (1)–(3) are affected by two kinds of basic physical variables: seal geometry, e.g., seal inside and outside diameters and conning angle; and operational variables, such as sealed fluid (lubricant) viscosity, sealed pressure, shaft speed, and clearance. The results of this research are applied to the seal test rig (Fig. 1) with fixed seal size and sealed pressurized water. In particular, the effects of two parameters are actually tested on the rig and are presented here. The first is the clearance—an operating variable, which has been sought as the variable to be adjusted. The second is the seal coning angle—a geometrical variable that always exists in practical mechanical seals, which strongly affects the rotordynamic coefficients and, therefore, the clearance. The effects of other parameters have been studied elsewhere [18].

In Fig. 2 the maximum normalized relative misalignment, $\gamma^*_{\max}$, is plotted as a function of the coning angle for several clearances at different sealed water pressures and shaft speeds. In order to avoid seal face contact it is necessary that the maximum normalized relative misalignment should always be less than $\gamma^*_{\text{critical}}$. In the present application $\gamma^*_{\text{critical}}$ is 1.25 (shown as a bold horizontal line), such that $\gamma^*_{\max}$ less than 1.25 indicates noncontacting operation, while any $\gamma^*_{\max}$ greater than 1.25 indicates face contact.

At a given shaft speed of 80 rad/s, Fig. 2(*a*) shows that when the sealed water pressure is relatively small (100 kPa), noncontacting operation can be attained only for a seal having small coning angles (<0.7 mrad) and small clearances (<2 μ). Fig. 2(*b*) shows that at the said speed noncontacting operation is possible at larger coning angles (2–10 mrad), but only at larger clearances (>4 μ) and only if the sealed water pressure is increased to 900 kPa. As shown in Figs. 2(*c*) and 2(*d*), higher shaft speeds contribute to widening of the ranges for viable coning and clearance, and it is attributed to an amplified aligning gyroscopic effect. For example, at low sealed pressure (100 kPa) but high shaft speed (360 rad/s) a seal having a coning angle of almost 2 mrad can be maintained in noncontacting mode of operation provided the clearance is smaller than 2μ. Similarly, at a high pressure (900 kPa) and a high speed (360 rad/s) a seal having coning angles in the range of 1–10 mrad can be maintained in noncontacting mode of operation provided that the clearance is 4 μ or larger. Gener-

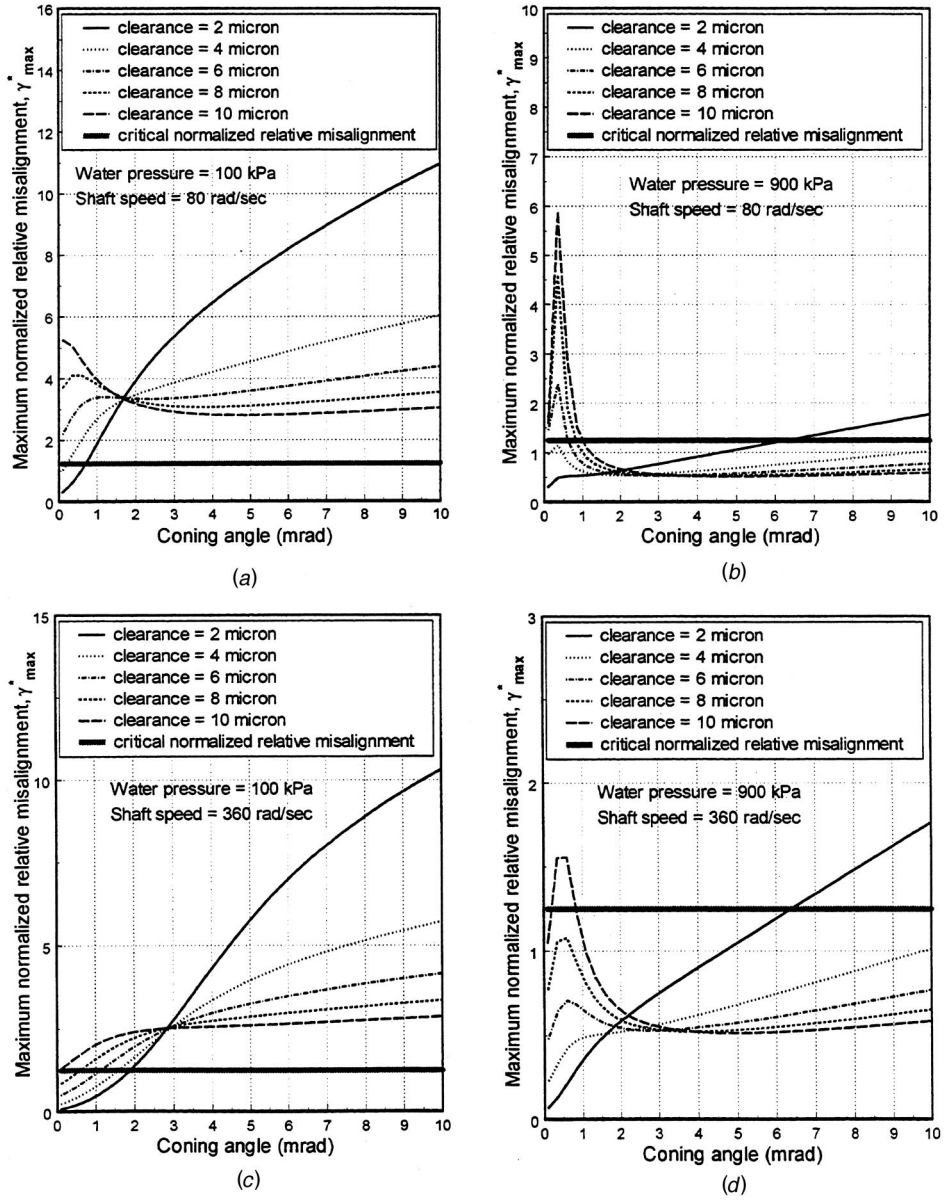


Fig. 2 Maximum normalized relative misalignment—coning angle relationship

ally, seals operating at small clearances and relatively small coning angles are preferred because they assume larger fluid film stiffness and damping, and yield smaller leakage. Operating seals with small coning angles and large clearances, or seals having large coning angles and small clearances should be avoided because they leak more and are more prone to contact.

Examination of all four cases presented in Fig. 2 reveals that if the seal under investigation possesses coning angles smaller than 1.5 mrad, then reducing the clearance invariably reduces the maximum normalized relative misalignment, and with it the likelihood of contact. And vise versa for large coning angle (>3 mrad), where reducing the clearance increases the maximum normalized relative misalignment.

This parametric analysis has been instrumental in establishing that small clearances and small coning angles are preferred for the safer operation of the current seal. As in any commercial application the current seal faces have been lapped to a flatness of about 1 μm maximum peak to valley height. Seal coning, however, is the outcome of various deformations caused by the hydraulic (sealed) pressure, centrifugal effects, and interface (sealing dam) heating. These effects are normally coupled and, therefore, in reality the coning is uncontrollable over a sustained period of time. The present study indicates, however, that in the current seal it is sufficient for the coning to fall into the range of ''small coning angles'' to allow for face contact to be eliminated via clearance adjustment (in this case, clearance reduction). From a controllability point of view, it was found [20] that the maximum normalized relative misalignment is most sensitive to clearance changes at a coning angle of about 0.5 mrad, or at coning angles that are greater than 3 mrad. At these coning angles the control effort is minimized. The best control results can be achieved, however, at the smallest coning angle, i.e., the critical coning angle needed for positive fluid film stiffness [19]. The critical coning angle varies with the clearance, and for the tested seal it ranges from 0.0984 mrad to 0.492 mrad for clearances of 2 μm to 10 μm, respectively.

As mentioned, stator and initial rotor misalignments are responsible for $\gamma_{\max}$. The parametric study performed above (that is based on a linearized analysis with the assumption of ''small perturbations'' about equilibrium) is very valuable for seal design

and performance prediction. It can also provide guidelines for action should contact occur. But other factors, such as uncertainties about the kinematics of the flexible support and its rotordynamic coefficients, machine deterioration, transients in sealed pressure or shaft speed, or unexpected shaft vibration, will all affect the dynamic behavior of the seal and, hence, the relative position between rotor and stator. This is where strict reliance on precise analysis loses effectiveness because some or all of the assumptions imbedded in the analysis may be borne out physically. Particularly, when intermittent face contact occurs the assumption and analysis of a ''noncontacting'' seal become irrelevant. The only analysis of intermittent seal face contact has been performed by Lee and Green [9] who observed higher harmonic oscillations (HHO), and offered a contact model based on a Fourier series expansion. Therefore, when it comes to actual diagnostics a phenomenological approach for contact detection is more appropriate, and indeed such an approach is adopted here.

Contact Elimination

Seal Test Rig. The noncontacting FMR mechanical face seal test rig used in this study (Fig. 1) is essentially identical to that in Lee and Green [9–11], equipped, however, with the more advanced real-time data acquisition and analysis system introduced by Zou and Green [12,13]. Other significant modifications include now the stator, which is made entirely of carbon graphite, and the rotor, which is made entirely of AISI 440C stainless steel. Both have been fabricated and lapped to industry standards by seal manufacturers. All these modifications facilitate more reliable measurement and determination of the relative position between rotor and stator.

The rotor is flexibly mounted on the rotating shaft through an elastomer O-ring. This allows the rotor to track the stator misalignment and to move axially. The seal stator assembly is composed of several components: the carbon stator, the spacer, and the stator holders. This design is capable of mechanically deforming the stator and produces seals with different coning angles [10]. For stability it is mandatory for the seal to maintain a converging gap in the direction of radial flow. For an outside pressurized seal the minimum seal film thickness has to be on the ID [19]. Transients in deformations caused by thermal effects, for example, occur at much slower pace than the time scale of interest in seal dynamics. Therefore, the two processes can be regarded as decoupled. For this reason the coning in the present test rig is induced by deforming the faces in the stator fixture and held fixed throughout the experiments. In that regard, data sufficient for dynamic analysis and monitoring is acquired in a fraction of a second, a time scale insignificant for any thermal deformations to occur. Moreover, as mentioned contact is detected here on a phenomenological basis, so it is not necessary to know where contact occurs; all that matters is that it occurs somewhere.

The stator assembly is fixed in the housing, which is made of three parts for convenience in machining, maintenance, and adjustment of the test rig. All possible leakage paths are sealed by O-rings. The sealed fluid in the housing is pressurized water. The shaft is connected to a spindle driven by a DC motor through two timing pulleys and a timing belt. A speed controller controls the motor speed.

Pressurized air is supplied from the main air supply line to the rotor chamber through holes in the housing and the shaft. It is sealed by a lip seal at one end and separated from the water by a contacting seal at the other end. The seal operates at an equilibrium clearance where the opening and closing forces are balanced. Changing the closing force by adjusting the air pressure in the rotor chamber (whether manually or by the computer through a voltage to pressure converter) varies the clearance. Further details of the test rig components, data acquisition and analysis can be found in the aforementioned references.

Clearance Adjustment. Three eddy current proximity probes mounted on the end of the housing measure the instantaneous dynamic response of the rotor along the shaft axis. These proximity probes have a bandwidth of about 10 kHz. They can measure the static and dynamic distances between their tips and the rotor face end surface. Key parameters including stator misalignment, rotor misalignment, relative misalignment between the rotor and the stator are calculated on-line in real-time from the probe measurements [12]. The clearance of the seal is obtained as follows: when the shaft is stationary, high air pressure is applied in the rotor chamber to ensure that the rotor is pressed against the stator, at which state the probe readings are taken. The average readings of the two probes mounted 180 deg apart represents the reference for zero clearance. The clearance of the seal at any time is then the difference between the instantaneous average readings of these two probes and the zero reference. (Sehnal et al. [5], and Etsion and Constantinescu [6], have made similar attempts to determine the clearance from proximity probe readings, but they reverted to indirectly estimating the clearance from a simplified equation applied to the measured leakage.) A low pass filter with a cut-off frequency of 1 kHz is used to eliminate high frequency cross-talk noise among the probes and also to serve as an anti-aliasing filter. The proximity probe signals in terms of the reduced voltages are sent through the A/D converters to a floating-point Digital Signal Processor (DSP). This DSP, supplemented by a set of on-board peripherals, such as analog to digital (A/D) and digital to analog (D/A) converters, comprise a universal board mounted in a personal computer. The seal clearance is then calculated and compared with the desired value. Based on the error signal between the measured clearance and the desired one a control signal is sent to the electropneumatic transducer (through the D/A converter), which provides a proportional air pressure signal. Zou and Green [13] provided an in-depth description and implementation of the said control scheme.

Experimental Results. The feasibility of contact elimination through clearance adjustment is now experimentally investigated. The proximity probe measurements are used as feedback to calculate the clearance. The experiments are conducted at 344.8 kPa water pressure and 28 Hz (176 rad/s) shaft speed, i.e., well within the range investigated in Fig. 2. The coning angle is set to 1 mrad, roughly twice as large as the largest critical value and sufficiently close to the optimal value of 0.5 mrad for minimum effort control (see above). Two sets of experiments are presented. Fig. 3 to Fig. 5 show the results for experiments with small stator initial rotor misalignments, both of 0.5 mrad, and Fig. 6 to Fig. 8 for experi-

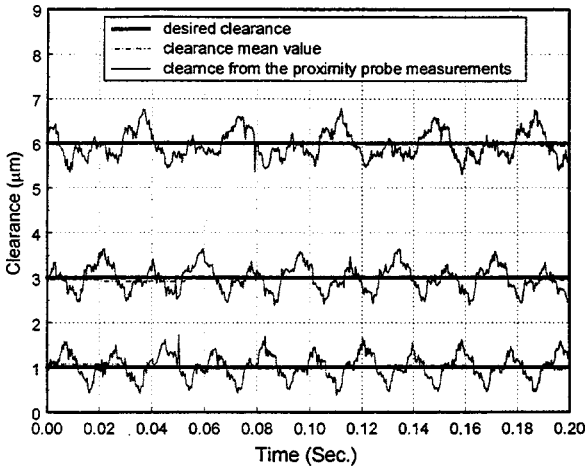
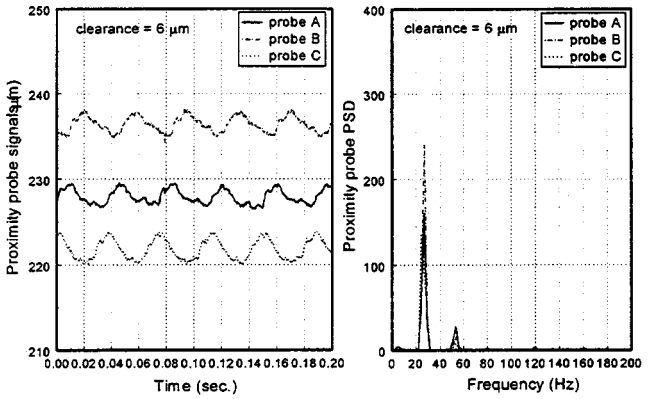


Fig. 3 Desired clearances and clearances calculated from proximity probe measurements (small stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 0.5$ mrad)



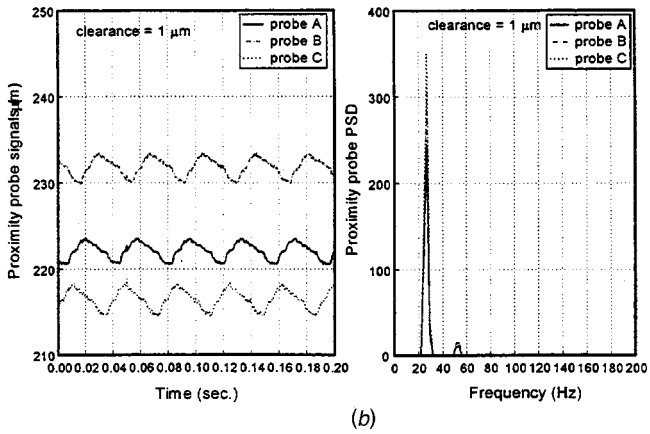


Fig. 4 Proximity probe signals and their PSDs for different clearances (6 μm and 1 μm) (small stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 0.5$ mrad)

mental results with large stator and initial rotor misalignments, both of 1.5 mrad. Other experiments for mixed misalignments [21] show similar behavior.

The desired clearance is stepwise set to 6 μm, 5 μm, 4 μm, 3 μm, 2 μm, and 1 μm, and held constant by a PI controller [13], to investigate one at a time the clearance effect on seal face contact. Figures 3 and 6 show the desired clearances, the actual clearances

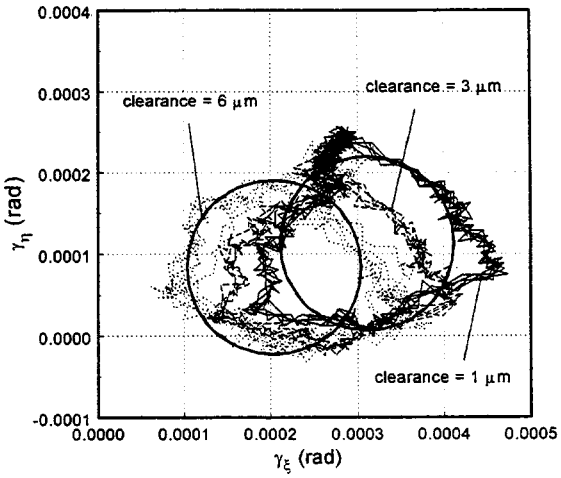


Fig. 5 Rotor angular misalignment orbit for different clearances (6 μm, 3 μm, and 1 μm) (small stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 0.5$ mrad)

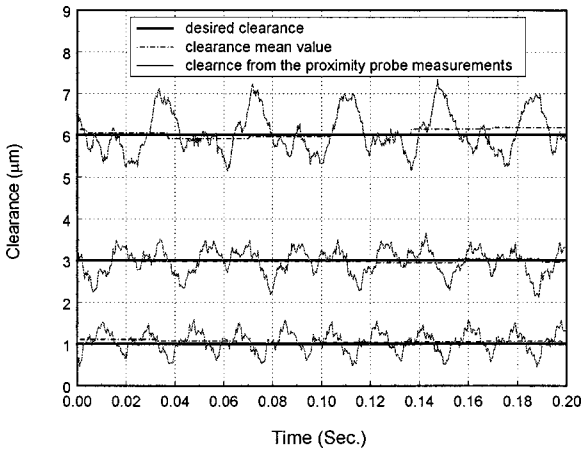
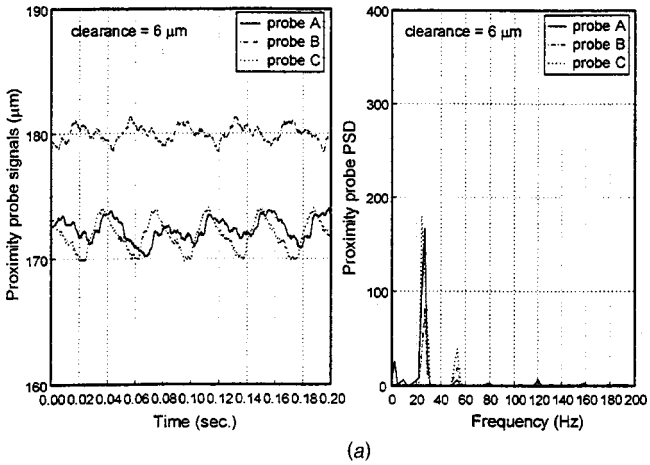


Fig. 6 Desired clearances and clearances calculated from proximity probe measurements (large stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 1.5$ mrad)

calculated from the measurements of the two proximity probes mounted 180 deg apart, and the calculated clearance mean value for each revolution (for clarity the graphs depict results for 6, 3, and 1 μm only). Unlike the theoretical assumption of constant clearance, the measured clearance is periodic and contains higher harmonic oscillations (HHO). The peak to peak value of the mea-



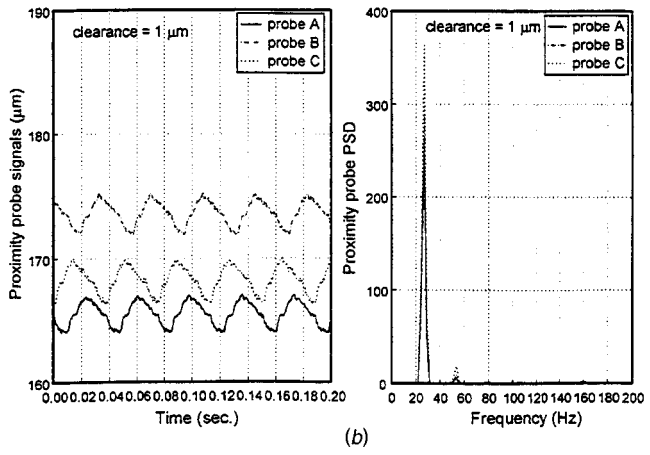


Fig. 7 Proximity probe signals and their PSDs for different clearances (6 μm and 1 μm) (large stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 1.5$ mrad)

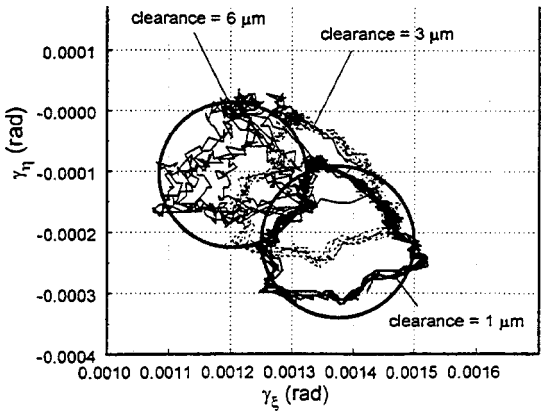


Fig. 8 Rotor angular misalignment orbit for different clearances (large stator and initial rotor misalignments $\gamma_s = \gamma_{ri} = 1.5$ mrad)

sured clearance decreases and there are less HHO present in the measured clearance signal as the clearance is reduced.

Because of the factors mentioned above and because the measured clearance oscillates, face contact is not determined based on the theoretical (or calculated) normalized relative misalignment. Instead, it is based on the pattern of the three probe signals and their power spectrum densities (PSD). Figures 4 and 7 show the three proximity probe signals and their associated power spectrum densities for desired clearances of 6 μm, and 1 μm, respectively. (The plots for all intermediate values show similar behavior, see Zou [21].) The PSD for all clearances have second higher harmonic oscillations (HHO), which are equal to twice the shaft rotating frequency. Certain levels of these HHO are inherent in the system. They are present even when the rotor runs without the stator in place. Other system components, particularly the O-ring flexible support, are suspected to induce these second HHO. It should be noted, however, that their energy level at the largest clearance case (6 μm) is much higher than that at the smallest clearance case (1 μm). Indeed, it is clearly visible (Figs. 4(*a*), and 7(*a*)) that when the clearance is large (6 μm), the shape and the peak to peak value of each probe signal is different from the others and is accompanied by the above mentioned pronounced HHO. As seal clearances decrease the shape and peak to peak value of the probe signals change as well, and tend to become similar to each other. When clearance reaches 1 μm the three probe signals are almost identical, where HHO have practically disappeared. It can also be noticed that in the largest clearance, Figs. 4(*a*) and 7(*a*) contain sub harmonics and super high harmonics. However, as the seal clearance decreases (particularly to 1 μm—see Figs. 4(*b*) and 7(*b*)), the sub and super higher harmonics have vanished. The absence of HHO in the PSD is an indication that noncontacting operation has been restored (see [9]).

The angular response orbit [12] is used as another means of diagnostics that capture the rotor angular dynamic behavior. It represents the locus of the magnitude of the rotor misalignment vector positioned at the instantaneous precession angle. Figures 5 and 8 show the rotor angular misalignment orbit for the desired clearances of 6 μm, 3 μm, and 1 μm. (The results of 5 μm, 4 μm, and 2 μm, are similar and were omitted for clarity.) Two circles are drawn in Figures 5 and 8 through orbit data obtained for the 6 μm clearance and for the 1 μm clearance (no circle is drawn through the 3 μm data because it is very similar to the 1 μm case). It is readily seen that the deviation of the data from a perfect circle in the 6 μm case is far bigger than that of the 1 μm cases. An odd shape cluster of data points in the 6 μm case of Fig. 5 (the blackened area shown at the top of the postulated circles), denotes part of the 6 μm orbit which clearly does not belong to the circle. Indeed, the orbit shape becomes more circular as the clearance decreases, which is another indication of noncontacting operation [12]. Therefore, the absence of high harmonics in the PSD plot for these data and the more circular orbits obtained for decreased clearance both show that contact can indeed be eliminated by such clearance adjustment. In summary, using the proximity probes' measurements as feedback in a straightforward control loop, interpreting the existence of high harmonics or non-circular orbits as contact, and applying the necessary closing force on the FMR to adjust the clearance, eliminates that contact. As mentioned, additional experimental results for the mixed cases of large stator misalignment with small initial rotor misalignment and vise versa also support these findings.

Conclusions

This research suggests that contact elimination may be achieved through clearance adjustment. The feasibility of doing so is experimentally investigated using FMR seal dynamics with large and small stator and rotor misalignments. Contact detection and elimination is demonstrated for two combinations of stator misalignments and initial rotor misalignment when both are 0.5 mrad, and then when both are 1.5 mrad. The clearance is calculated in real-time from the measurements of proximity probes and is varied through a pneumatic closing force adjusting mechanism in the test rig. All experimental results show that contact can be eliminated through clearance adjustment. Moreover, these experimental results are predicted by the parametric study. Particularly for the tested seal and contrary to intuition, it is found that the maximum normalized relative misalignment decreases with the clearance and, therefore, the likelihood of face contact decreases as well. Therefore, once seal contact is detected by the phenomena described (i.e., HHO detected by PSD and/or noncircular orbit plots), it can be eliminated by adjusting the clearance to a smaller value.

Worthy of notice is that most previous research on controlled mechanical seals used temperature control instead of clearance control. When the seal clearance reduces, the temperature at the sealing dam may rise because of increased viscous heating. In such a case, the action that would be taken by a controlled system relying on temperature would be to increase the clearance. This, contrary to the finding of this research, is likely to have detrimental consequences on the seal.

Likewise it must be highlighted that the clearance cannot be decreased just without bounds. Clearly if the clearance is decreased to the point where surface asperities start rubbing, the basic assumption of ideal noncontacting operation ceases to be accurate. The goal in seal design, and now in seal monitoring and control, is to maintain the smallest clearance possible with either minimal face contact or no face contact at all. It is believed that the technology developed in this research can assist in accomplishing this goal.

Advancing this technology in real practice is to automate the two processes of contact detection, and then contact elimination. This consequently will necessitate the development of a strategy for automatic clearance adjustment to a new desired value, which may be based upon a certain measure of the variations of PSD and/or orbit plots from a known or ideal noncontacting behavior.

Acknowledgments

The authors wish to express their appreciation to the Office of Naval Research for support of research grant N00014-95-1-0539, entitled *Integrated Diagnostics*. Dr. Peter Schmidt serves as program officer. The authors also wish to express their appreciation to Mr. Andrew Flaherty and Rexnord Corporation for machining the seal rotor, to Mr. Laurence Thorwart and Mr. David Erich of Pure Carbon Company for providing the carbon seals used in this study. This project is also supported in part by a Georgia Tech foundation Grant, E25-A77, made by Mr. Gilbert Bachman. This support is gratefully acknowledged.

References

[1] Buck, Jr., G. S., 1980, ''A Methodology for Design and Application of Mechanical Seals,'' ASLE Trans., **23**, No. 3, pp. 24–252.

[2] Will, Jr., T. P., 1982, ''Experimental Observation of a Face-Contact Mechanical Shaft Seal Operation on Water,'' Lubr. Eng., **38**, No. 12, pp. 767–772.

[3] Etsion, I., and Burton, R. A., 1979, ''Observation of Self-Excited Wobble in Face Seals,'' ASME J. Lubr. Technol., **101**, No. 4, pp. 526–528.

[4] Metcalfe, R., 1982, ''Dynamic Whirl in Well-Aligned, Liquid-Lubricated End-Face Seals with Hydrostatic Tilt Instability,'' ASLE Trans., **25**, No. 1, pp. 1–6.

[5] Sehnal, J., Sedy, J., Zobens, A., and Etsion, I., 1983, ''Performance of the Coned-Face End Seal with Regard to Energy Conservation,'' ASLE Trans., **26**, No. 4, pp. 415–429.

[6] Etsion, I., and Constantinescu, I., 1984, ''Experimental Observation of the Dynamic Behavior of Noncontacting Conedface Mechanical Seals,'' ASLE Trans., **27**, No. 3, pp. 263–270.

[7] Green, I., 1989, ''Gyroscopic and Support Effects on the Steady-State Response of a Noncontacting Flexibly-Mounted Rotor Mechanical face Seal,'' ASME J. Tribol., **111**, pp. 200–208.

[8] Green, I., 1990, ''Gyroscopic and Damping Effects on the Stability of a Noncontacting Flexibly-Mounted Rotor Mechanical Face Seal,'' *Dynamics of Rotating Machinery*, Hemisphere Publishing, Bristol, PA, pp. 153–173.

[9] Lee, A. S., and Green, I., 1994, ''Higher Harmonic Oscillation in a Flexibly Mounted Rotor Mechanical Seal Test Rig,'' ASME J. Vibr. Acoust., **116**, No. 2, pp. 161–167.

[10] Lee, A. S., and Green, I., 1995, ''Physical Modeling and Data Analysis of the Dynamic Response of Flexibly Mounted Rotor Mechanical Seal,'' ASME J. Tribol., **117**, No. 1, pp. 130–135.

[11] Lee, A. S., and Green, I., 1995, ''An Experimental Investigation of the Steady-State Response of a Noncontacting Flexibly Mounted Rotor Mechanical Face Seal,'' Trans. ASME, J. Tribol., **177**, No. 1, pp. 153–159.

[12] Zou, M., and Green, I., 1997, ''Real-Time Condition Monitoring of Mechanical Face Seal,'' *Proceedings the 24th Leeds-Lyon Symposium on Tribology*, London, Imperial College, pp. 423–430.

[13] Zou, M., and Green, I., 1999, ''Clearance Control of a Mechanical Face Seal,'' STLE Tribol. Trans., **42**, No. 3, pp. 535–540.

[14] Salant, R. F., Miller, A. L., Kay, P. L., Kozlowski, J., Kay, W. E., and Algrain, M. C., 1987, ''Development of an Electrically Controlled Mechanical Seal,'' *Proceedings 11th International Conference on Fluid Sealing*, BHRA, pp. 576–595.

[15] Heilala, A. J., and Kangasneimi, A., 1987, ''Adjustment and Control of a Mechanical Seal Against Dry Running and Severe Wear,'' *Proceedings 11th International Conference on Fluid Sealing*, BHRA, pp. 548–575.

[16] Etsion, I., Palmor, Z. J., and Harari, N., 1991, ''Feasibility Study of a Controlled Mechanical Seal,'' Lubr. Eng., **47**, No. 8, pp. 621–625.

[17] Wolff, P., and Salant, R. F., 1995, ''Electronically Controlled Mechanical Seal for Aerospace Applications: Part II—Transient Tests,'' Tribol. Trans., **38**, No. 1, pp. 51–56.

[18] Zou, M., Dayan, J., and Green, I., 1999, ''Parametric Analysis for Contact Control of a Noncontacting Mechanical Face Seal,'' *Proceedings of Vibration, Noise & Structural Dynamics*, Venice Italy, April 28–30, 1999, pp. 493–499.

[19] Green, I., 1987, ''The Rotor Dynamic Coefficients of Coned-Face Mechanical Seals with Inward or Outward Flow,'' ASME J. Tribol., **109**, No. 1, pp. 129–135.

[20] Dayan, J., Zou, M., and Green, I., 1999, ''Sensitivity Analysis for the Design and Operation of a Noncontacting Mechanical Face Seal,'' J. Mech. Eng. Sci., in print.

[21] Zou, M., 1998, ''Real-Time Monitoring and Control of a Noncontacting Mechanical Face Seal,'' Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.

Studies of Spray Breakup and Mixture Stratification in a Gasoline Direct Injection Engine Using KIVA-3V

Dennis N. Assanis
e-mail: assanis@engin.umich.edu

Sang Jin Hong

Akihiro Nishimura

George Papageorgakis[1]

Bruno Vanzieleghem

W. E. Lay Automotive Laboratory,
University of Michigan,
1231 Beal Avenue,
Ann Arbor, MI 48109-2121

The Low Pressure spray Breakup (LPB) model of Papageorgakis and Assanis has been implemented in the multi-dimensional code KIVA-3V as an alternative to the standard Taylor Analogy Breakup (TAB) model. Comparison of spray predictions with measurements shows that the LPB model, in conjunction with the standard k-ε turbulence model, has the potential for simulating the evolution of hollow cone sprays with acceptable fidelity, both from qualitative and quantitative standpoints. After validating the LPB model, illustrative studies of mixture stratification are conducted for a Direct Injection Gasoline (DIG) combustion chamber resembling the Mitsubishi design. The effects of reverse tumble strength and injection timing on mixture quality in the vicinity of the spark plug are explored. Overall, the study demonstrates how the KIVA-3V code with the LPB model can contribute to the optimization and control of mixing in DIG engines.
[S0742-4795(00)00303-3]

Introduction

DIG spark-ignition engines can achieve significantly reduced specific fuel consumption and CO_2 emissions, while meeting other emission standards. Charge stratification and control of mixture distribution near the spark plug permit operation with very lean air-fuel ratios. Lean burning is thermodynamically attractive and reduces heat transfer losses, thus improving fuel efficiency. Control of the load by regulating the amount of fuel injected into the chamber can dramatically reduce pumping losses and improve transient response. Direct injection in and fuel vaporization in the cylinder improves volumetric efficiency due to charge cooling [1]. Charge stratification and cooling enables the use of higher compression ratios without knock, thus further enhancing thermal efficiency.

Many automotive manufacturers are responding to the quest for fuel economy by developing DIG systems (e.g., [2–6]). With advances in the capabilities of multi-dimensional engine computer models, it has become increasingly possible to use them in DIG studies of mixture preparation and combustion under a wide range of engine operating conditions. For instance, Han et al. [7–9] used the multi-dimensional reactive flow code KIVA-3 [10] to study the effect of intake flow structures and injection timing on fuel-air mixing in a DIG engine with a flat top piston. Naitoh et al. [11] performed numerical studies on a DIG engine using axial swirl to create sufficient mixing. Duclos et al. [12] performed KIVA-3 studies of intake, injection and combustion in a DIG engine under both homogeneous and stratified operating conditions. The above studies have demonstrated that the fuel-air mixing process is significantly affected by the flow structures, with swirl and tumble providing a more uniform fuel-air distribution. It has also been shown that fuel injection timing directly affects spray tip penetration, spray-wall impingement and fuel-air mixing. While such numerical studies can significantly contribute to the understanding of fundamental DIG issues, the accuracy of their predictions critically depends on the fidelity of the underlying models for injection, breakup and turbulence.

Several models of engine spray atomization have been proposed in the context of CFD computations, but none seems to adequately capture all breakup mechanisms (vibrational, bag, multimode, shear, catastrophic) that happen under different Weber numbers. Adjustable constants are typically used in all breakup models to bring predictions in line with experimental data under different conditions. For the KIVA family of codes, the TAB model [13,10] and the surface wave growth model [14] have been extensively used to simulate breakup of relatively high injection pressure, solid cone diesel sprays. However, the applicability of such models to low injection pressure, hollow cone sprays, typical of DIG engines is questionable. Since breakup directly influences the spray structure and fuel air mixing, it is essential to develop and apply higher fidelity breakup models in DIG computations.

Miyamoto and Kobayashi [15] used a sheet atomization model for air-assisted hollow-cone injectors based on a sheet stability argument. The sheet thickness was computed by solving the coupled liquid and gas governing equations inside the injectors, but secondary drop breakup was not considered. Lee and Bracco [16] also used a sheet atomization model for the breakup model of DIG engines. Lagrangian equations were used to solve for the motion of an intact sheet outside the injector, but the computed liquid sheet and subsequent drops were sensitive to grid settings. Han et al. [17] used a variant of the TAB model to predict hollow-cone sprays by replacing TAB's χ^2 distribution (which is based on diesel spray measurements and thus provides an overpopulation of large size drops) with a Rosin-Rammler distribution. Although their model improved predictions of spray motion, tip penetration and shape compared to the original TAB model, aerodynamic effects such as air entrainment were still not well predicted.

Papageorgakis and Assanis [18] developed an alternative Low Pressure Breakup (LPB) model, applicable between the Rayleigh and first wind-induced regimes, that accounts for both aerodynamic and surface tension effects. In this study, their LPB model is implemented in the KIVA-3V code and its potential for capturing the evolution of hollow-cone sprays is compared against the TAB model in the light of experimental data. After demonstrating the promise of the LPB model, illustrative fuel-air mixing studies

[1]Dr. George Papageorgakis is currently with Exxon Production Research, Rm C-330, P.O.B. 2189, Houston, TX 77252-2189.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division December 12, 1998; final revision received by the ASME Headquarters January 7, 2000. Technical Editor: H. Nelson.

are conducted for a combustion chamber resembling the Mitsubishi design. The effects of reverse tumble strength and injection timing on mixture quality in the vicinity of the spark plug are explored. Overall, the paper demonstrates how KIVA-3V studies with a realistic breakup model can contribute to the optimization and control of mixing in DIG engines.

Spray Breakup Modeling

The Taylor Analogy Breakup (TAB) model [13] is standard in the KIVA-3V code. The model is based on an analogy between an oscillating and distorting drop and a damped, forced harmonic oscillator. The model considers the effects of the liquid viscosity on the oscillations of small droplets, and predicts that there is no unique critical Weber number of droplet breakup. Nevertheless, the TAB model can only track one oscillation mode, whereas in reality there are many such modes. Several investigators have shown that the TAB model is best suited for spray calculations under high injection pressures where aerodynamic effects are predominant, such as with diesel engines [14].

The LPB model developed by Papageorgakis and Assanis [19] has shown promise in capturing the large-ligament droplet breakups that occur due to Rayleigh-Taylor instabilities within the droplet flow field. The LPB model is based on the fundamental work of Harper et al. [20] who solved the distortion problem of a single droplet moving at constant acceleration. The analysis addressed the interior flow within the droplet and its interaction with the exterior domain. A force balance was conducted for the distorted and accelerating droplet by considering the pressure force due to the interior liquid field, the exterior aerodynamic forces and surface tension. To account for the aerodynamic effect at low Weber numbers and for a surrounding gas density much smaller than the droplet density, the external pressure on the droplet surface was assumed to be a steady axisymmetric distribution of arbitrary form. Through coupling of the assumed external pressure with the linear solution for the internal flow field within the droplet, the non-dimensional disturbance of the droplet surface can be expressed as

$$\overline{\eta}^{(1)} = \sum_{2}^{\infty} \frac{n(2n+1)C_n}{4\beta_n^*} P_n(\chi)[\cos(\beta_n^* t^*) - 1]$$

$$\overline{\eta} = \frac{\eta}{r_0}, \quad t^* = \frac{tU_\infty}{r_0}, \quad \beta_n^* = \frac{[n(n-1)(n+2)]^{1/2}}{\mathrm{We}^*}, \quad \mathrm{We}^* = \frac{\rho_l}{\rho_g}\mathrm{We}.$$

In implementing the LPB model in KIVA-3V, we postulated that breakup occurs when the non-dimensional radial displacement of the distorted droplet in the flow direction exceeds 5 percent of the radius of the undistorted spherical droplet. This suggests that further distortion will be driven by surface tension effects. Furthermore, only the first two terms of the equation are considered, as they represent the first two fundamental modes of the droplet's distortion. Even to that approximation, the equation turns out to be the superposition of two harmonic distortions with frequencies that are not integer multiplicatives of each other. Since kinetic energy cannot be determined as readily as for the linear oscillator of the TAB model, calculation of the Sauter mean radius is no longer based on conservation of energy in the droplet before and after breakup. Rather, it is assumed that a droplet can break up into three, four, or five droplets with the same probability according to a random number generator. The droplet radii after breakup are calculated based on conservation of mass. Droplet radii are then used to determine the Sauter mean diameter of the distribution. It is further assumed that whenever a droplet breaks up in an odd number of droplets (three or five), one of the resulting droplets moves in the axial direction, whereas the rest move in the direction prescribed by the initial spray cone angle. Lastly, the breakup time is determined using Newton-Raphson's method, as it is impossible to solve the equation and obtain a closed form solution with respect to time.

Model Validation

To validate the LPB breakup model, its ability to predict quantitatively and qualitatively the evolution and size distribution of the hollow cone spray was compared to both the TAB model and experimental measurements. Both turbulence models available in the KIVA-3V code, i.e., the standard k-ε model [21] modified to take into account volumetric expansion effects as well as spray-turbulence interactions, and the Renormalized Group (RNG) based k-ε model [22] were explored in conjunction with each of the breakup models. The main difference between these two turbulence models is in the constants used in the equation for the turbulent kinetic energy dissipation rate.

The experimental study of DIG sprays by Zhao et al. [23] was used to assess the KIVA-3V predictions. The global spray development process was visualized using a two-dimensional laser Mie scattering technique. The tests were conducted at room temperature inside a constant-volume chamber whose ambient pressure could be varied to simulate the spray development process at different fuel injection timings. The Sauter Mean Diameter (SMD) of the droplet distribution, spray tip penetration, and velocity components of the droplets were measured under different fuel injection pressures and timings.

The injector used was a hollow cone injector, with a cone angle of 65 deg and a spray angle of 5 deg. The nozzle area was calculated to be 0.0005 cm^2. These data were derived from the spray pictures and SMD data at the nozzle exit. The chamber that was simulated had a diameter of 7 cm and a length of 12 cm. This domain was found to be sufficient in size to model the spray without wall interaction. Calculations with the two alternative breakup models (TAB and LPB) and the two alternative turbulence models were performed on increasingly dense grids to establish grid independence. Figure 1 shows the average total velocity distribution for increasingly dense grids. The runs were executed using the LPB breakup model in conjunction with the k-ε turbulence model. The grid spacing between nodes was 1 cm for the coarse grid, 0.5 cm for the intermediate grid, and 0.33 cm for the dense grid. Comparisons between SMD predictions and data were made at a plane located 3 cm downstream from the injector. The difference between the intermediate and dense grid velocity distributions was found to be small, thus allowing us to assume that grid independence was achieved with the dense grid. Very similar results were found using the other available combinations of spray breakup and turbulence models. The maximum difference between the dense and intermediate grid predictions

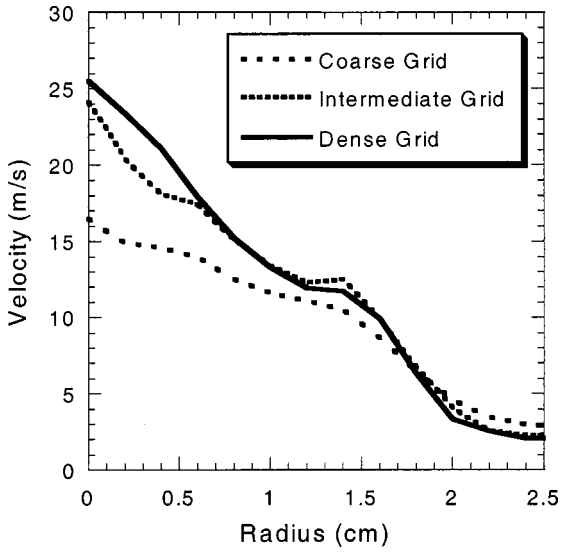


Fig. 1 Average total velocity distribution computed with LPB and k-ε models for three increasingly dense grids

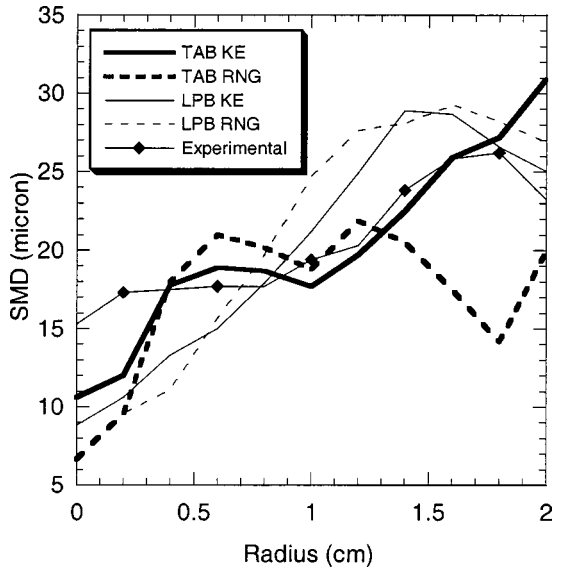


Fig. 2 SMD comparison with experimental data for combinations of LPB and TAB with k-ε and RNG turbulence models

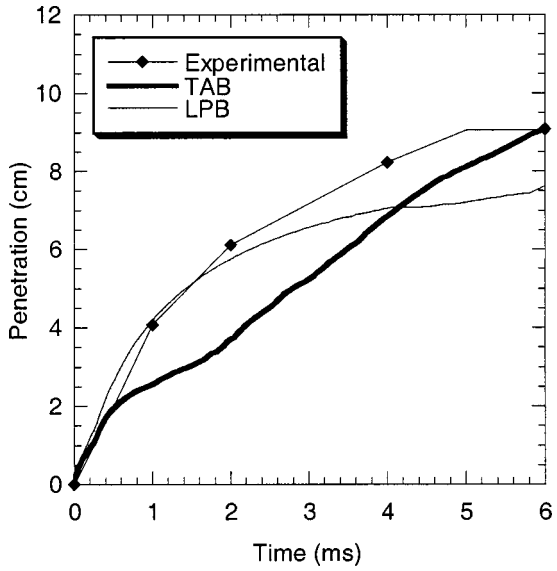


Fig. 3 Comparison of spray tip penetration computed using the TAB and LPB models with experimental data [23]

was found to be approximately 12 percent, and occurred close to the axis of the spray. Differences between the predictions of the two latter grids were significantly smaller away from the axis, where in fact the majority of the droplets were found.

Predictions of computed spray SMD using the various model permutations are compared to the experimental measurements in Fig. 2. It can be observed that the RNG turbulence model results in an erratic, wavy prediction of the SMD, especially in combination with the TAB model. While the RNG model can improve predictions in some recirculatory flows, such as the backward facing step, in a recent study, Papageorgakis and Assanis [19] found that the standard k-ε model can predict shear flows, such as confined coflow-jets more accurately than the RNG model. Based on our similar finding here, all results shown below are obtained with the standard k-ε model for turbulence. When comparing the LPB model with the TAB model, both in combination with the k-ε turbulence model, it can be observed that both models significantly underpredict droplet sizes close to the injector axis. However, it must be noted that only a small number of droplets are present close to the injector axis and it is often difficult to achieve accurate measurements at this location. While the TAB model compares more favorably with the experimental data for radii between 0.5 and 1.5 cm, the trend of the experimental results is not well matched at large radii where the number of droplets is larger. On the other hand, the LPB model overpredicts droplet sizes for radii between 0.5 and 1.5 cm, but compares favorably with experimental data in both magnitude and trend for larger radii, where a larger number of droplets are present. Overall, the SMD comparison alone can not conclusively show which of the two models is superior.

In order to assess further the accuracy of the LPB model, computed spray tip penetrations using both breakup models were compared against the measurements reported by Zhao et al. [23]. The data shown in Fig. 3 compare the square root of the sum of the squares of the axial and radial tip penetrations. Clearly, the LPB model displays a much more physical behavior than the TAB model. In the early stages of injection, the LPB tip penetration does exhibit the linear variation with time that is associated with the intact core. Later on, LPB spray tip penetration displays the experimentally observed square root behavior with time. In contrast, the TAB model predicts an almost linear increase in penetration with time. This suggests that the LPB model has potential for higher fidelity predictions than TAB.

Figure 4 presents a visualization of spray evolution obtained from KIVA-3V using the TAB and LPB breakup models, both in combination with the k-ε turbulence model. In addition, the corresponding spray image from the experimental study by Zhao et al. [23] is shown. The background color in the modeling images represents the air velocity in the chamber, while round points represent the spray parcels. The differences between the TAB and LPB model predictions are significant, with the LPB model creating a well defined hollow cone spray. By comparing the spray images from the experimental study of Zhao et al. [23] to the images predicted by KIVA using the alternative breakup models, we conclude that the LPB model can track the spray development process more realistically than the TAB model. In particular, the LPB model captures the experimentally observed recirculation zones at the periphery of the spray at the later stages of injection, or after the end of injection. Nevertheless, the aerodynamic effects are over-estimated in some regions, i.e., more ambient gas seems to be entrained into the spray compared to the experimental results. This suggests that alternative formulations of turbulence need to be investigated in connection with the LPB model.

Mixture Stratification Study

DIG engines use a stratified-charge lean-burn, late injection strategy at part load, which favors high fuel economy. At high load (above an equivalence ratio $\phi = 0.5$), almost all reported designs have adopted early fuel injection occurring during the induction stroke to give adequate timing for preparing a close to stoichiometric mixture with little stratification and to optimize power. Early injection, during the latter part of the inlet stroke, improves volumetric efficiency due to charge cooling form the latent heat of vaporization of the fuel. This study explores the potential for achieving mixture stratification under full load ($\phi = 1.0$) and low speed (1000 rpm) using late injection. This strategy could be used as an intermediate step for transitioning from low speed, low load stratified operation, to low speed, high load homogeneous operation when accelerating.

Grid and Computational Strategy. A representative DIG combustion chamber with a spherical compact cavity in the piston surface was considered for our study. The bore and stroke of the engine modeled were 8.2 cm and 8.9 cm, respectively. The compression ratio was 12. A computational grid was constructed using almost 39,000 cells at BDC. The grid and other important geo-

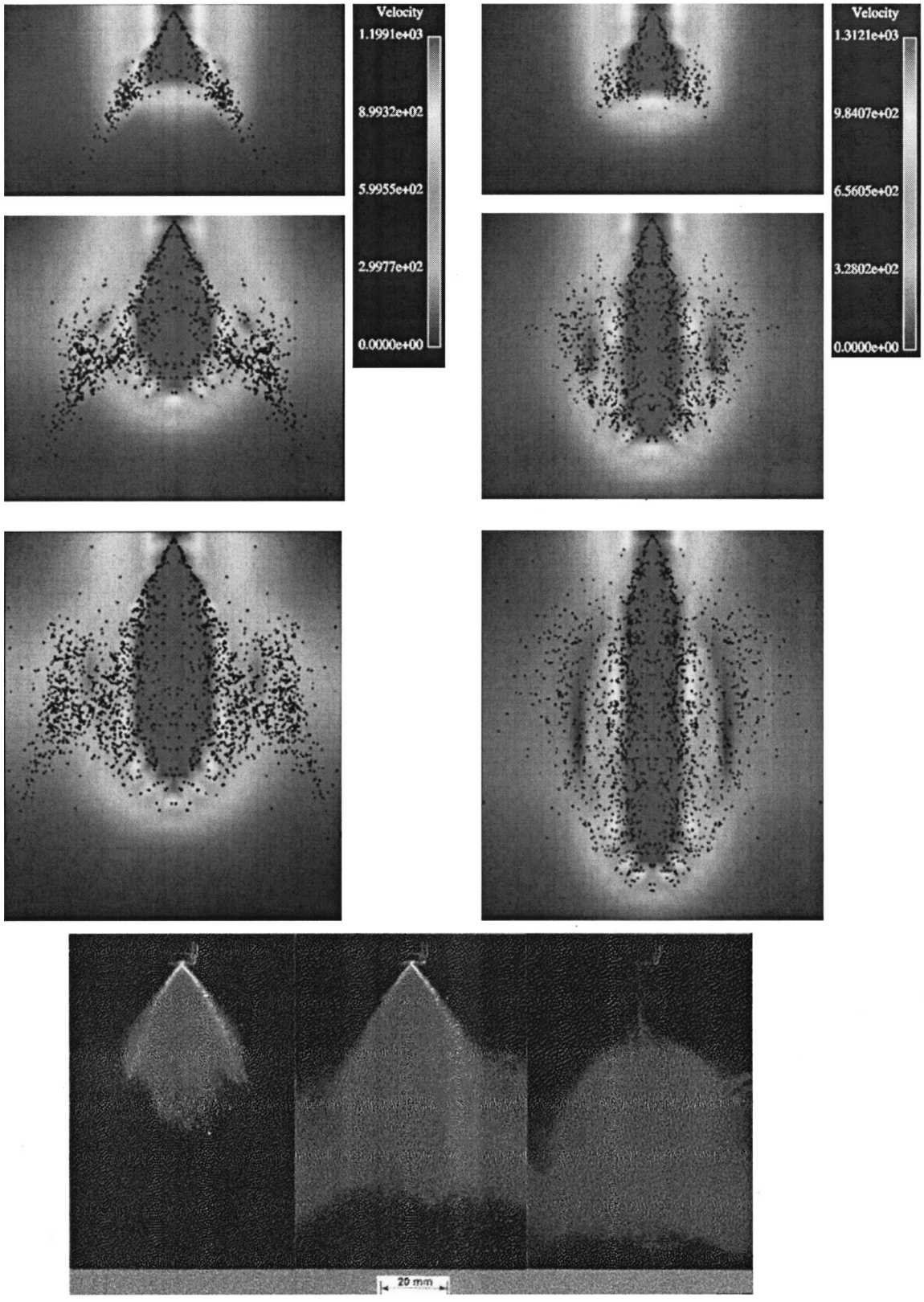


Fig. 4 Visualization of computed spray droplet distribution using LPB model (left) and TAB model (right) every 2 ms starting at $t=2$ ms. Experimental spray images [23] at 1, 4, and 6 ms for the modeled conditions

metric aspects related to the placement and targeting of the injector and spark plug are shown in Fig. 5. The injection angle was set at 45 degrees. Injection velocity and injection duration were set at 50 m/s and 10 deg CA, respectively.

In a previous study, Duclos et al. [12] performed a numerical simulation for a Mitsubishi style combustion chamber, including modeling the intake and exhaust processes. Their study indicated that two tumble vortices are set up by this four-valve chamber

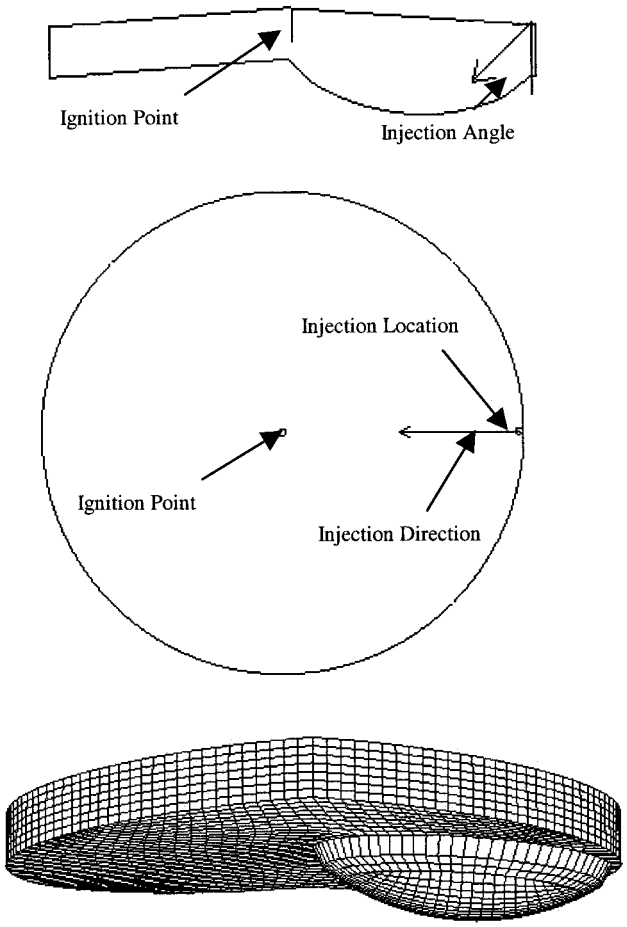


Fig. 5 Vertical and horizontal views of combustion chamber, and perspective view of the grid at TDC

design during the induction process. One of the tumble vortices is located below the exhaust valve, while the other one is centered over the bowl of the asymmetric, curved-top piston. The tumble motion below the exhaust valve was shown to be dissipated during the compression stroke. In contrast, the clockwise tumble motion centered over the piston bowl was shown to dominate the fuel air mixing process. Based on these findings, we have elected to focus our study on how the strength of the latter tumble vortex affects mixture stratification. It will be referred to as reverse tumble vortex since it is in the opposite sense than the tumble produced in standard port fuel-injected, four-valve, flat top piston engine designs.

In an actual DIG engine, the strength of the reverse tumble vortex would depend on the intake port and valve design, number and placement of valves, valve actuation and control strategy, and engine operating conditions. In order to decouple our investigation from the details of the intake system design, as well as to expedite the numerical simulation, runs were conducted for the closed part of the cycle only. While the reverse tumble vortex prescribed in this manner would be somewhat different than the port-generated one in an actual engine, this approach offers attractive flexibility and is still very useful for conceptual studies. An initial reverse tumble motion is imposed around an axis perpendicular to a vertical cross-section at the beginning of a simulation (start of compression). The reverse tumble is imposed in the same way the swirl is imposed in KIVA-3V. The tumble velocities increase as a Bessel function of the distance from the tumble center, and are set to zero at the wall. The strength of the reverse tumble is indicated by the tumble ratio, defined as the ratio of the angular speed of the air over that of the crankshaft.

Illustration of the Mixing Strategy. In this study, the fuel is injected late in compression and directed toward the spherical piston cavity. Subsequently, the spray is redirected towards the spark plug by the reverse tumble and the impingement action on the cavity surface. The spacing between the fuel injector and the spark plug is intentionally wide to provide additional time for fuel breakup, evaporation and mixing with air. Thereby, an air-fuel mixture of desired mixture strength can be prepared around the spark plug by the moment of spark firing. While Kume et al. [2] have demonstrated this injection strategy for equivalence ratios up to 0.5 (low loads), we will explore whether it is possible to use this strategy with full load operation ($\phi = 1.0$).

The fluid mechanics associated with late fuel injection are illustrated for a baseline case (CASE A) in Fig. 6, which tracks flow and mixing histories at different instants during compression. The prescribed reverse tumble ratio is 1.0 at the start of compres-

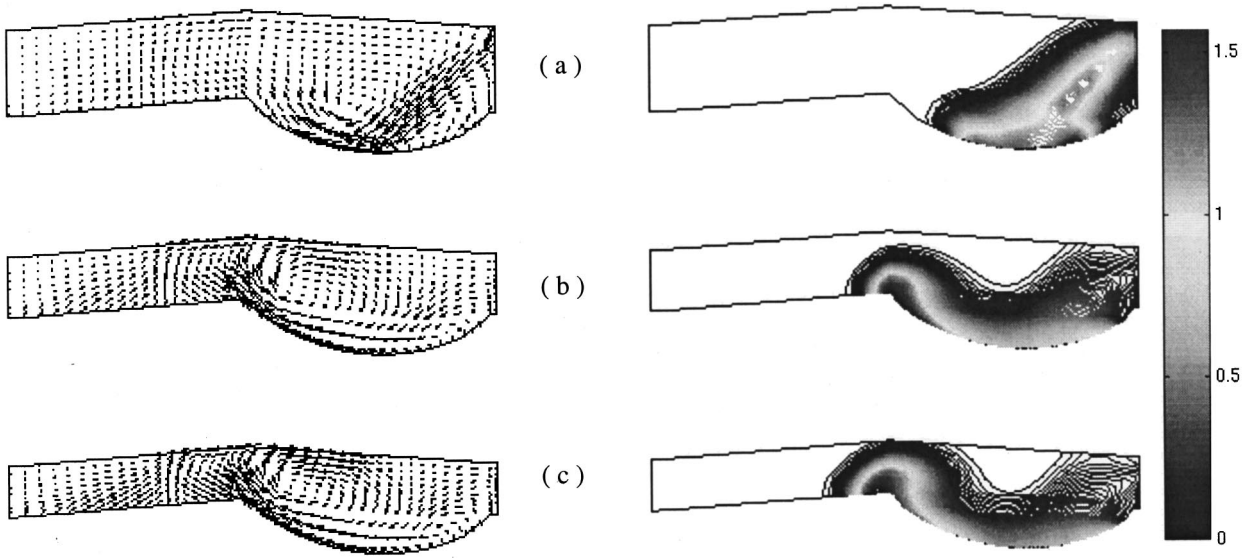


Fig. 6 Consecutive snapshots of velocity profiles (left) and equivalence ratio contours (right) shown at (*a*) 30 deg BTDC, (*b*) 20 deg BTDC, and (*c*) 150 deg BTDC. Case A: baseline configuration with injection timing=40 deg BTDC and tumble ratio=1.0.

sion. Gasoline fuel (modeled as C_8H_{17} in KIVA-3V) is injected at 40 deg Before Top Dead Center (BTDC). The left hand side of Fig. 6 shows the velocity vectors, and the right hand side shows the evolution of the local fuel-air equivalence ratio as the piston moves up towards TDC of compression. The spray has a basically hollow cone structure. An intense airflow is generated at the center of the hollow cone due to the movement of the fuel spray. The velocity vectors of Fig. 6(*a*) clearly show the turbulent air motion imposed by the fuel injection. As the downward airflow penetrates into the cavity, it changes into an upward flow after striking the walls. The gas motion created by this piston design intensifies the initially imposed reverse tumble motion, and contributes to the roll-up vortex that carries the fuel vapor to the vicinity of the spark plug (Figs. 6(*b*) and 6(*c*)). As illustrated by Fig. 6(*c*), this tumble vortex is contained in the right half of the combustion chamber, above the piston cavity. Its spread outside that region is attenuated by the strong squish flow that is developing as the piston is moving up from the left hand side of the combustion chamber, thus displacing fluid radially inward, from the exhaust side towards the intake side.

The above fluid mechanics have a direct impact on the fuel transport and mixing with the air. Additional critical factors are fuel evaporation and impingement on the piston cavity which are modeled with the standard models available in KIVA-3V. In the equivalence ratio plots of Fig. 6(*a*), the stratification of the fuel-air mixture is evident. Fuel starts to evaporate in areas of intense turbulence along the injection direction, and initially accumulates over the piston cavity, thus producing the higher equivalence ratio spots (with ϕ ranging between 1.45 and 1.8) along the bowl. The outside contours of the spray correspond to an equivalence ratio of almost zero. The central rich core can be seen to be dragged along the piston cavity by the tumble motion in Fig. 6(*b*). The roll-up vortex subsequently transports the concentrated mixture towards the spark plug, achieving an equivalence ratio of 0.9 in its vicinity at the time of spark, 15 deg BTDC. Note that the V-neck shape indicated by the iso-concentration profiles is attributed to the centrifugal action of the reverse tumble and the squish flow from the left side of the combustion chamber. The geometric confinement as the mixture strikes the firedeck surface and starts to spread is evident.

This baseline case suggests that late injection strategies can be considered for full load operation at low engine speeds, as the latter can provide adequate time for fuel evaporation. For our case (1000 rpm), 60 percent of the fuel had evaporated by the spark timing, i.e., in a crank angle window of 25 deg. Further validation work of the evaporation and wall impingement models is recommended, however, in order to increase our confidence in numerical assessments of the merits of late injection strategies.

Effect of Reverse Tumble Strength on Charge Stratification. The impact of the strength of the intake-generated reverse tumble vortex on charge stratification is explored here via parametric studies for higher and lower tumble ratios than the baseline. Figure 7 shows velocity vectors and equivalence ratio contours at spark timing (15 deg BTDC) for a tumble ratio=0 (Case B) and a tumble ratio=2 (Case C); for both cases B and C, the injection timing was kept at 40 deg BTDC (same as for Case A). When we compare the flow patterns at the time of spark for the baseline CASE A (tumble ratio=1, Fig. 6(*c*)) with Case B and Case C, it is evident that they all exhibit similar general characteristics. As expected, the reverse flow becomes more organized and centered in the bowl as initial tumble ratio increases. The differences in flow patterns appear to be greater as tumble ratio is increased from weak to baseline, than from baseline to strong. In all cases, the acceleration of the flow in the vicinity of the edge of the piston bowl is evident. Clearly, the piston cavity is instrumental in inducing the reverse tumble motion that is necessary to drag the fuel along the piston cavity. Once the high velocity zone at the edge of the bowl is reached, fuel diffuses away from the wall, towards the spark plug. The strength of the initially imposed reverse tumble vortex, however, does affect the time of arrival of the roll-up vortex at the spark plug. Case A appears to have an optimum tumble strength to transport the fuel to the spark plug by the set ignition time. In Case B, the tumble motion is of insufficient strength to transport the rich fuel core to the spark plug by the time of spark. For Case C, too much tumble has caused the fuel to move beyond the spark plug, because the high velocity zone at the edge of the bowl was reached earlier in the cycle.

Figure 8 presents the evolution (by crank-angle) of the equivalence ratio near the spark plug as tumble ratio is varied from 0 to 2. Injection timing was fixed at 40 deg BTDC for all cases. For tumble ratios of either 1 or 2, the mixture concentration in the vicinity of the spark plug is within flammability limits. Nevertheless, for tumble ratio=1, both a relatively higher equivalence ratio is reached at the plug by spark timing, and the mixture does not get leaned out as rapidly as for tumble ratio=2, following the spark; this would reduce the chance of bulk quenching. It is clear that for zero tumble strength, there would be no fuel near the

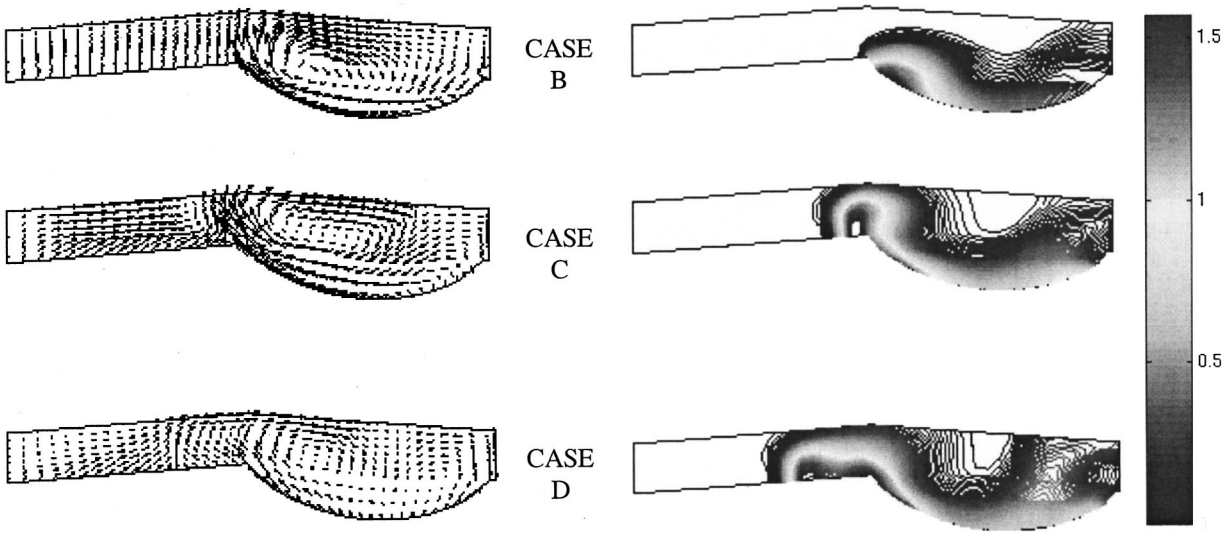


Fig. 7 Comparison of velocity distribution (left) and equivalence ratio contours (right) at the point if ignition (15 deg BTDC) for different tumble strengths and injection timings. Case B: tumble ratio=0, injection at 40 deg BTDC; Case C: tumble ratio=2, injection at 40 deg BTDC; Case D: tumble ratio=1.0, injection at 60 deg BTDC).

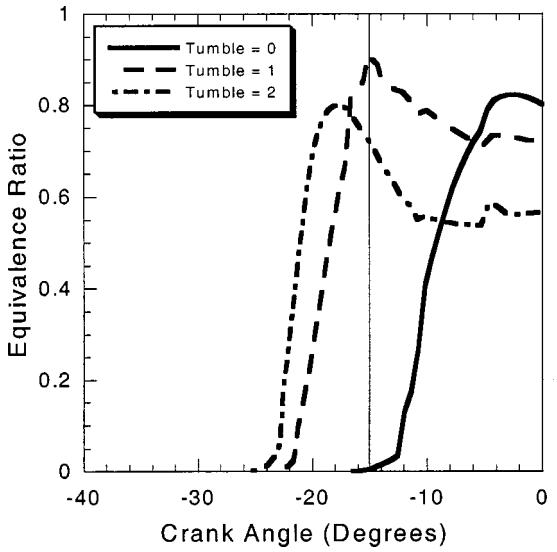


Fig. 8 Equivalence ratio near the spark plug for tumble ratio varied from 0 to 2

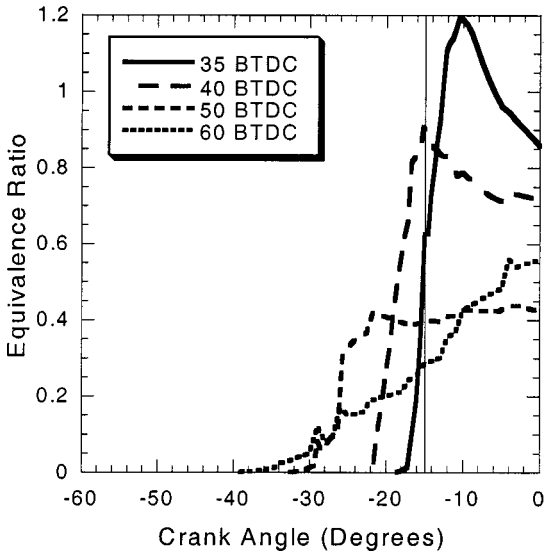


Fig. 9 Equivalence ratio near the spark plug for varied injection timings

spark plug had the spark been fired at 15 deg BTDC. This is because the fuel reaches the high velocity zone at the edge of the bowl late in the cycle, due to the low tumble strength. In this case, the lack of externally imposed tumble could be somewhat compensated by retarding the spark to 3 deg BTDC. While this would lead to acceptable ignition conditions, it would tend to slow down burn rates during expansion.

Our results indicate that a well designed piston cavity, along with a properly targeted injection system, can in their own right generate reverse tumble motion. As a result, fuel air mixture within flammability limits will eventually be transported to the vicinity of the spark plug, even with no initial tumble at BDC. However, in order for the flammable mixture to arrive to the plug at an optimum time (i.e., not too late in compression), a certain level of intake flow generated tumble is desirable. Note that the peak equivalence ratios for tumble strengths varying from 1 to 2 are relatively close together in both timing and magnitude. Above a certain tumble ratio, it appears that there is little to be gained by the extra tumble strength, and in fact it could lead to overleaning problems.

Effect of Injection Timing on Charge Stratification. At any given operating condition, the injection timing needs to be selected strategically so as to provide sufficient time to promote spray evaporation and diffusion for suppression of over-rich or over-lean regions around the plug. The fluid mechanics associated with earlier versus later injection have been investigated by comparing distributions for an injection timing of 60 deg BTDC (Case D, Fig. 7) to those of the baseline case with an injection timing of 40 deg BTDC. The tumble ratio=1.0 for both cases. With the injection timing set at 40 deg BTDC, an ignitable mixture is transported near the spark plug by 15 deg BTDC. For injection at 60 deg BTDC, however, the equivalence ratio contours show that a significant amount of fuel penetrates beyond the piston cavity. This results in overleaning of the fuel, creating a mixture that cannot be ignited. This behavior can be explained in the light of the fact that injection timing affects both the location where the penetrating spray impinges on the piston cavity, and the subsequent spreading of the mixture. With early injection, the spray penetrates faster as the density of the charge air is lower and the radially inward squish motion is not intensified yet.

Figure 9 shows the evolution (by crank angle) of the equivalence ratio near the spark plug, as injection timing is varied from 35 deg to 60 deg BTDC, at a constant tumble ratio of 1.0. Clearly, achieving proper mixture strength by the time of spark firing is very sensitive to injection timing. For a tumble ratio=1.0, an injection timing of 40 deg BTDC gives a readily ignitable mixture (ϕ close to 0.9) near the spark plug. The earlier injection cases (50 deg and 60 deg BTDC) result in overleaning, for reasons discussed above. The problem appears to be major in the sense that even altering spark timing cannot compensate it. On the other hand, a later injection timing (35 deg BTDC) can produce a readily ignitable mixture in the vicinity of the spark plug (ϕ =1.2), but spark timing must also be retarded by 5 deg (to 10 deg BTDC). However, this selection might eventually slow down burning, and thus not produce maximum brake torque at that speed.

Conclusions

A low-pressure breakup (LPB) model has been implemented in the KIVA-3V code. Comparison of predictions with experimental measurements has indicated that the LPB model has the potential for predicting the evolution of DIG hollow cone sprays with higher fidelity than the standard TAB model, both from quantitative and qualitative standpoints. From this study, it also appears that the proposed LPB model works better in conjunction with the standard k-ε model. Subsequently, the modified KIVA-3V code with the LPB model was applied to the study of mixture stratification in a realistic DIG combustion chamber. The effects of injection timing and reverse tumble strength have been investigated. The following conclusions have been drawn:

- An intense airflow is generated at the center of the hollow cone spray as a result of momentum transfer from the spray to the air. As the downward airflow penetrates into the piston cavity, it changes into an upward flow after striking the walls. The gas motion created by this piston design intensifies the initially imposed reverse tumble motion, and contributes to the roll-up vortex that carries the fuel vapor to the vicinity of the spark plug.
- A well designed piston cavity, along with a properly targeted injection system, can in their own right generate reverse tumble motion. As a result, fuel air mixture within flammability limits will eventually be transported to the vicinity of the spark plug, even with no initial tumble at BDC. However, for the flammable mixture to arrive to the plug at an optimum time (i.e., not too late in compression), a certain level of intake flow generated tumble is required.

• For a given reverse tumble strength, injection timing needs to be optimized such that there is sufficient, but not excessive, time allowed for evaporation and mixing. When fuel is injected too early, spray overpenetration in a lower charge density, lower squish environment leads to overleaning.

• For a DIG engine with a properly optimized intake system and piston design, it appears that late injection strategies can be viable for full load operation, at least at low engine speeds, provided that the injection timing is optimally set.

Acknowledgments

This study has been supported in part by the Honda Initiation Grant Program administered by Honda R&D North America. The authors gratefully acknowledge Dr. Fu-Quan Zhao of Daimler-Chrysler for providing the experimental photographs of the sprays and valuable comments.

Nomenclature

n = index
P_n = n^{th} zonal harmonic
r = droplet radius
t = time
U_∞ = free stream velocity
We = Weber number
η = amplitude of surface distortion
ρ = density

Subscripts and Superscripts

l = liquid
g = gas
o = undisturbed dropelt
(1) = first order approximation

References

[1] Anderson, R. W., et al., 1996, ''Understanding the Thermodynamics of Direct Injection Spark Ignition Combustion System: An Analytical and Experimental Investigation,'' SAE Paper 962018.
[2] Kume, T., Iwamoto, Y., Iida, K., Murakami, M., Akishino, K., and Ando, A., 1996, ''Combustion Control Technologies for Direct Injection SI Engine,'' SAE Paper 960600.
[3] Iwamoto, Y., Noma, K., Nakayama, O., Yamauchi, T., and Ando, H., 1997, ''Development of Gasoline Direct Injection Engine,'' SAE Paper 970541.
[4] Kuwahara, K., Ueda, K., and Ando, H., 1998, ''Mixing Control Strategy for Engine Performance Improvement in a Gasoline Direct Injection Engine,'' SAE 980158.
[5] Harada, J., Tomita, T., Mizuno, H., Mashiki, Z., and Ito, Y., 1997, ''Development of Direct Injection Gasoline Engine,'' SAE Paper 970540.
[6] Tomoda, T., Sasaki, S., Sawada, D., Saito, A., and Sami, H., 1997, ''Development of Direct Injection Gasoline Engine—Study of Stratified Mixture Formation,'' SAE Paper 970539.
[7] Han, Z., Reitz, R. D., Claybaker, P. J., Rutland, C. J., Yang, J., and Anderson, R. W., 1996, ''Modeling the Effects of Intake Flow Structures on Fuel/Air Mixing in a Direct-Injected Spark-Ignition Engine,'' SAE Paper 961192.
[8] Han, Z., Reitz, R. D., Yang, J., and Anderson, R. W., 1997, ''Effects of Injection Timing on Air-Fuel Mixing in a Direct-Injection Spark-Ignition Engine,'' SAE Paper 970625.
[9] Han, Z., Fan, L., and Reitz, R. D., 1997, ''Multidimensional Modeling of Spray Atomization and Air-Fuel Mixing in Direct-Injection Spark-Ignition Engine,'' SAE Paper 970884.
[10] Amsden, A. A., 1993, ''KIVA-3: A KIVA Program with Block-Structured Mesh for Complex Geometries,'' Los Alamos National Lab., LA-12503-MS.
[11] Naitoh, K., et al., 1998, ''Numerical Optimization of the Fuel Mixing Process in a Direct-Injection Gasoline Engine,'' SAE 981440.
[12] Duclos, J. M., and Zolver, M., 1998, ''3D Modeling of Intake, Injection and Combustion in a DI-SI engine under Homogeneous and Stratified Operating Conditions,'' COMODIA 98, pp. 335–340.
[13] O'Rourke, P. J., and Amsden, A. A., 1987, ''The TAB Method for Numerical Calculation of Spray Droplet Breakup,'' SAE Paper 872089.
[14] Reitz, R. D., 1987, ''Modeling Atomization Processes in High-Pressure Vaporizing Sprays,'' Atomization Spray Technol., **3**, pp. 309–337.
[15] Miyamoto, T., and Kobayashi, T., 1996, ''Structure of Sprays from an Air-Assist Hollow-Cone Injector,'' SAE 960771.
[16] Lee, C. F., and Bracco, F. V., 1995, ''Comparisons of Computed and Measured Hollow-Cone Sprays in an Engine,'' SAE 950284.
[17] Han, Z., Parrish, S., Farrel, P. V., and Reitz, R. D., 1997, ''Modeling Atomization Processes of Pressure-Swirl Hollow-Cone Fuel Sprays,'' Atomization Sprays, **7**, pp. 663–684.
[18] Papageorgakis, G. C., and Assanis, D. N., 1996, ''A Spray Breakup Model for Low Injection Pressures,'' Int. Commun. Heat Mass Transfer, **23**, No. 1, pp. 1–10.
[19] Papageorgakis, G. C., and Assanis, D. N., 1999, ''Comparison of Linear and Non-Linear RNG-Based k-ε models for Incompressible Turbulent Flows,'' Numer. Heat Transfer, Part B, **35**, pp. 1–22.
[20] Harper, E. Y., Grube, G. W., and Chang, I. D., 1972, J. Fluid Mech., **52**, No. 3, pp. 565–591.
[21] Launder, B. E., and Spalding, D. B., 1972, *Mathematical Models of Turbulence*, Academic Press, New York.
[22] Han, Z., and Reitz, R. D., 1995, ''Turbulence Modeling of Internal Combustion Engines Using RNG k-e Models,'' Combust. Sci. Technol., **106**, p. 207.
[23] Zhao, F., Yoo, J., Liu, Y., and Lai, M., 1996, ''Spray Dynamics of High Pressure Fuel Injectors for DI Gasoline Engines,'' SAE Paper 961925.

Experimental Investigation to Study Convective Mixing, Spatial Uniformity, and Cycle-to-Cycle Variation During the Intake Stroke in an IC Engine

Woong-Chul Choi
FloCoTec, Inc.,
Columbus, OH 43212
flocotec@cris.com

Yann G. Guezennec
The Ohio State University,
Department of Mechanical Engineering,
Columbus, OH 43210
guezennec.1@osu.edu

The work described in this paper focuses on experiments to quantify the initial fuel mixing and gross fuel distribution in the cylinder during the intake stroke and its relationship to the large-scale convective flow field. The experiments were carried out in a water analog engine simulation rig, and, hence, limited to the intake stroke. The same engine head configuration was used for the three-dimensional PTV flow field and the PLIF fuel concentration measurements. High-speed CCD cameras were used to record the time evolution of the dye convection and mixing with a 1/4 deg of crank angle resolution (and were also used for the three-dimensional PTV measurements). The captured sequences of images were digitally processed to correct for background light non-uniformity and other spurious effects. The results are finely resolved evolution of the dye concentration maps in the center tumble plane. The three-dimensional PTV measurements show that the flow is characterized by a strong tumble, as well as pairs of cross-tumble, counter-rotating eddies. The results clearly show the advection of a fuel-rich zone along the wall opposite to the intake valves and later along the piston crown. It also shows that strong out-of-plane motions further contribute to the cross-stream mixing to result in a relatively uniform concentration at BDC, albeit slightly stratified by the lean fluid entering the cylinder later in the intake stroke. In addition to obtaining phase-averaged concentration maps at various crank angles throughout the intake stroke, the same data set is processed for a large number of cycle to extract spatial statistics of the cycle-to-cycle variability and spatial non-uniformity of the concentration maps. The combination of the three-dimensional PTV and PLIF measurements provides a very detailed understanding of the advective mixing properties of the intake-generated flow field. [S0742-4795(00)00103-4]

Introduction

The experimental study of fuel distribution and fuel mixing in the cylinder is usually addressed by performing Planar Laser Induced Fluorescence (PLIF) measurements in engines, typically under motored conditions. In this case, the fuel is tagged with a dye which fluoresces when illuminated by a powerful laser at the proper wavelength. Details of the gaseous PLIF principles can be found in the work of Seitzman and Hanson [1] and others. Application of this technique to IC engine research can be found in recent works by Baritaud and Heinze [2], Wolff et al. [3], Zhao et al. [4], Neij et al. [5], and others. While the details of the implementation vary among the various researchers, the successful implementation of gaseous PLIF technique usually requires the use of a pulsed Nd:Yag laser (working in the UV band) inducing the short-lived fluorescence of molecular tracers such as bi-acetyl or acetone. Due to the very faint fluorescence signal and their short life, intensified, gated CCD camera (ICCD) is required to capture the fluorescence pattern in an engine suitably modified to allow optical access. While this technique has been very successful in the past decade, its implementation is expensive, lengthy and complex, which prevents its use as an engine development tool and restricts it to an engine research environment.

Meanwhile, the use of a Water Analog Engine Simulation Rig ([6–8], and others) has gained considerable importance to very effectively and practically study the flow fields generated during the intake stroke. The recent advent of whole field three-dimensional Particle Tracking Velocimetry (three-dimensional PTV) [9] now allows detailed, quantitative mapping of the complete three-dimensional flow field to be routinely carried out particularly at the end of the intake stroke. Hence, the technique promotes the understanding of the links among the complex head geometry, valve events and the resulting flow fields [10–13]. It is well accepted that the organized flow structures formed during the intake stroke contribute strongly to the subsequent combustion process, as the precursor of the flow seen at the time of spark [14]. In addition, they also contribute strongly to the initial mixing of the fuel and air in the cylinder through advective mixing.

In the broader field of experimental fluid mechanics and turbulence, PLIF using fluorescent dye in liquid flows has been extensively used in the past 20 years to perform qualitative flow visualization as well as a host of quantitative measurements ([15–18], and many others). The main advantage of using dye-based PLIF concentration measurements in liquids is the ease of use, as the fluorescence signal (usually induced by an Argon-Ion laser sheet illuminating a Fluorescein or Rhodamine-based dye) is very intense and long-lived. Furthermore, the intrinsic time scales (outer scales) of the liquid flows are considerably longer than in gases at comparable Reynolds numbers. This means that the imaging can be performed using commonly available lasers and cameras. One of the primary differences between concentration field measurements in liquid flows and concentration field measurements in gaseous flows is the large difference in Schmidt number. Dye in

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division June 1, 1999; final revision received by the ASME Headquarters October 21, 1999. Technical Editor: D. Assanis.

liquid is characterized by high Schmidt number, indicating that diffusive mechanisms are insignificant compared to convective effects, while diffusive effects are significant compared to convective mixing in gases. In the context of IC engines during the intake stroke, the overall fuel concentration in the cylinder is governed by the advective component due to the filling of the cylinder. The convective length scale of the process is directly linked to the piston motion, i.e., the fluid entering the cylinder during the intake stroke will convect a distance of the order of the following stroke length:

$$l_c \approx S,$$

where l_c is the convective length scale, and S is the stroke. On the other hand, the diffusive length scale can be estimated as (assuming a Schmidt number of order 1)

$$l_d \approx \sqrt{D_{AB}\tau} \approx \sqrt{\nu\tau},$$

where l_d is the diffusive length scale, D_{AB} is the molecular diffusivity between the two species, ν is the kinematic viscosity, and τ is the time scale of the process. However, in turbulent flows such as the ones encountered in the cylinder of IC engines, the diffusive process is not controlled by time-limited molecular diffusion (as represented in the equation above), but is rather strain-limited by the small turbulent eddy activity [18]. In this context, the diffusive length scale of relevance is given by

$$l_d \approx \sqrt{\nu/\varepsilon},$$

where ε is the strength of the strain rate field. From classical turbulence scaling, the strain rate field scales as follows:

$$\varepsilon \approx l/\lambda \approx S/\lambda \approx (\mathrm{Re})^{-1/2}.$$

During the intake stroke, the convective time scale of interest is the time it takes for the piston to move over a distance equal to the stroke

$$\tau_c = S/\overline{V}_p,$$

where $\overline{V}_p$ is the mean piston speed. In summary, the ratio of the length scale of the convective and diffusive processes in this turbulent flow can then be estimated as

$$l_d/l_c \approx (\mathrm{Re})^{-3/4}.$$

Under typical IC engine conditions, the Reynolds number (Re) is order of 3000 at idle to 20000 at maximum speed. Thus, the diffusive length scale is of the order 1 mm or less compared to a stroke of the order of 100 mm, typically. Under these conditions, the mixing is dominated by the convective process and the diffusion effects can be safely neglected, at least as far as the large scale fuel distribution in the cylinder is concerned. Therefore, the use of a high Schmidt number dye in a water analog engine simulation rig is justified for the study of the overall fuel distribution during the intake stroke. This approximation allows to replace complex gaseous PLIF measurements of fuel distribution with a relatively simple liquid-based PLIF to study the advective mixing process during the intake stroke. Furthermore, it also allows to perform spatially resolved three-dimensional velocity measurements using the three-dimensional PTV technique under the same conditions, hence providing a powerful tool to develop a mechanistic understanding of this mixing process.

Experiments

Three-Dimensional PTV Experimental Setup. The experiments were carried out in our water analog engine simulation rig. This rig consists of 1:1 scale transparent piston/cylinder assembly onto which an actual engine head is mounted. The piston and valves are actuated by servomotors to duplicate the actual kinematics of the simulated engine. Water is chosen as the working fluid to slow down the flow, while matching the actual Reynolds number encountered under actual engine conditions. The use of water naturally restricts the experiments to an incompressible simulation. Under the tested conditions, the flow during the intake stroke can be considered incompressible and therefore, the water analog simulation represents an accurate simulation of the flow. A more complete description of the water analog engine simulation rig can be found in Trigui et al. [14]. A schematic of the experimental setup is also shown in Fig. 1.

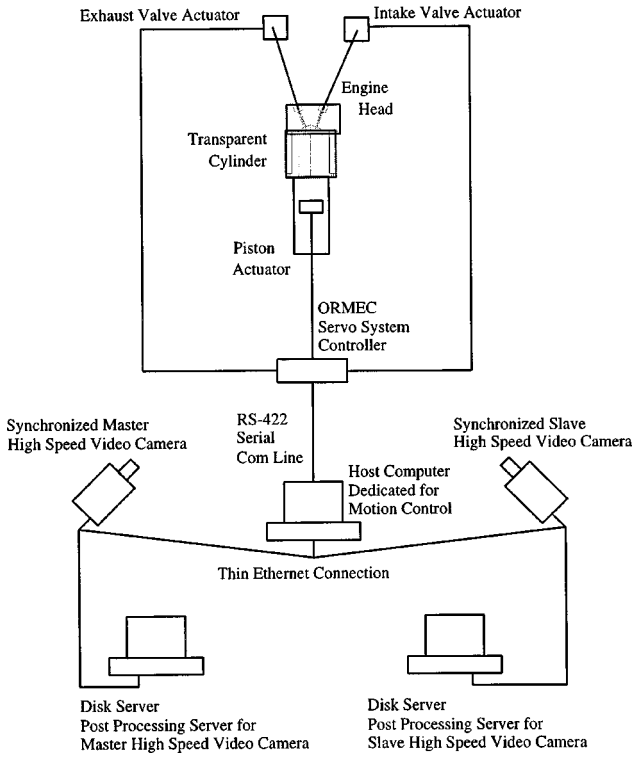


Fig. 1 Schematic of experimental setup for three-dimensional PTV and PLIF

The water is seeded with small, neutrally buoyant particles and is illuminated by a high intensity strobe light placed under the piston crown. A pair of high-speed digital cameras (Phantom II, with 256 MB of memory) is used to acquire sequences of images of the particles trajectories viewed from two nominally perpendicular directions. Dedicated synchronization electronics ensures that the motion controller, the strobe and the two cameras are properly synchronized.

Three-Dimensional Particle Tracking Velocimetry (Three-Dimensional PTV) Processing. Our state of the art implementation of the three-dimensional PTV technique [9] is based on a short sequence of 5 successive stereo image pairs (5 images per view) obtained from a pair of synchronized high speed video cameras. A brief description of the key features of the three-dimensional PTV technique will be recalled here for clarity. For more details, the reader is referred to the original paper. The short sequences of 5 consecutive images from each camera are preprocessed to enhance the image quality and then analyzed to determine two-dimensional positions of the particles in each frame and in each view. Based on the two-dimensional position data, the two-dimensional tracks in each view are automatically established based on the concept of path coherence [9]. Once the two-dimensional particle tracks are established, all sources of optical distortion during the image acquisition process and possible errors from the experimental setup are taken into account through an *in situ* calibration approach [19]. Normally the *in situ* calibration experiment is carried out prior to the actual experiment. Finally the three-dimensional stereo matching process is performed based on the calibrated two-dimensional tracks from each view. The

three-dimensional PTV results consist of randomly distributed three-dimensional raw velocity vectors over the entire volume of interest (in this case, within the cylinder). Typically, they are post-processed for extraction of related parameters and data presentation. In the current study, we used the commercial software FloTrack™, Version 2.1 developed by FloCoTec, Inc. which integrates all these concepts of three-dimensional PTV and *in situ* calibration techniques.

For the current three-dimensional PTV experiments, the dual high-speed digital cameras acquired the images at the rate of 288 frames per second (fps). This corresponds to 1/4 deg of crank angle per frame for the engine case simulated. This temporal resolution is required to adequately resolve the high speeds encountered in the flow field during the early part of the intake stroke. Furthermore, we acquired images over 100 engine cycles in order to achieve statistical convergence of the three-dimensional velocity data. Due to the high demand on the image storage space and the limited image buffer space available on the high-speed video cameras, we used two additional PCs with large hard disks networked with the digital cameras through a thin Ethernet connection. These computers were essentially the disk servers for the image acquisition system as well as the image pre- and post-processing servers.

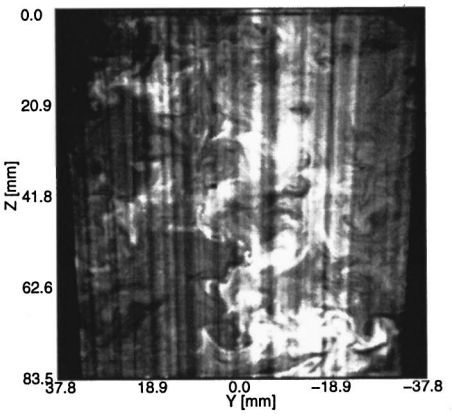
PLIF Experimental Setup. The same experimental setup described above was also used for the PLIF experiments, with a few minor modifications. The strobe placed under the piston for the PTV experiments was replaced with a 5W Spectra-Physics Argon-Ion laser with sheet-forming optics. Only one high-speed digital camera mounted perpendicular to the laser sheet was used to capture the PLIF sequences of images. The laser sheet was vertical in the tumble plane in the center of the cylinder. A fast acting solenoid valve was used to control the dye injection from a pressurized dye reservoir. The timing and duration of the injection is controlled by the piston and valve motion controller. The dye injector consists of a small hypodermic stainless steel tube mounted upstream of the intake valves at the same location and orientation where an actual fuel injector would normally be mounted for port fuel injection. Nonetheless, it should be realized that this dye injection is not to simulate complicated fuel injection process in an actual intake port of an IC engine where the phase change of the fuel as well as port wall wetting phenomena play important roles. The focus of the current experiment is to look at the convective mixing process closely related to the large scale flow motion further downstream of the intake generated flow field.

PLIF Image Processing. For the PLIF experiments, essentially a one-camera version of the above set-up was used. A total of 50 cycles were acquired for proper statistical convergence of the concentration statistics. Prior to acquiring the temporal evolution of the concentration field, a preliminary experiment was performed in order to acquire the ensemble averaged background images for each crank angle. To that effect, a uniform, pre-mixed solution of dye and water was illuminated by the laser sheet at all crank angles. The results from this uniform concentration imaging served to normalize the spatial non-uniformity of the laser sheet intensity (due to the Gaussian beam intensity profile, as well as various masking by the moving piston and rings). A more detailed description of the use of these ensemble averaged background images will be provided in the next paragraph.

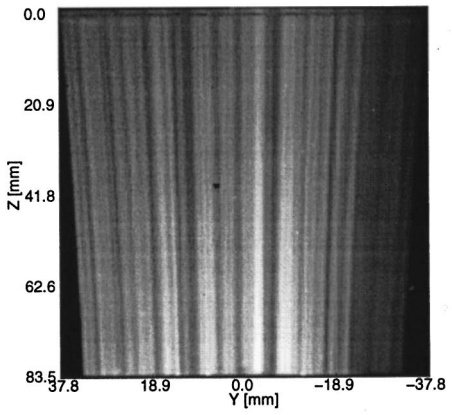
Images with a uniform, pre-mixed solution of dye and water corresponding to the average laser intensity at each crank angle were acquired. For better reliability of this normalization process, five such cycles were acquired and averaged at each crank angle:

$$\overline{I}_{\mathrm{cal}}(x,y)=\frac{1}{5}\sum_{i=1}^{5} I_{\mathrm{cal},i}(x,y),$$

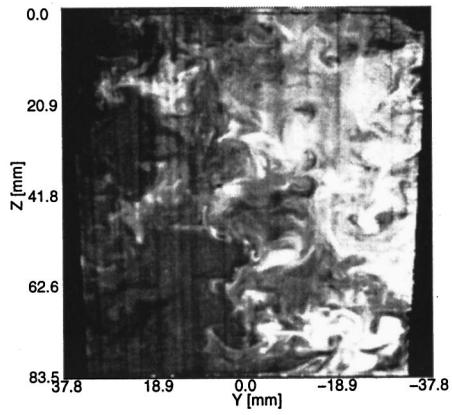
where $\overline{I}_{\mathrm{cal}}(x,y)$ is the ensemble averaged background image and $I_{\mathrm{cal},i}(x,y)$ is the ith background image. These images were used as background intensity maps to normalize each individual image from the PLIF experiments. Fig. 2(*a*) shows a raw (uncorrected) image for one realization at one crank angle of the PLIF imaging. Several features can be observed in this raw image (and in all other images). First, the light intensity is not uniform across the cylinder due to the Gaussian intensity profile of the laser beam used to form the illuminating light sheet. Second, line-like features are created by minute defects in the cylindrical lens and/or the piston crown, which effectively creates rays of lesser intensity. Obviously, these features are artifacts from the laser light illumi-



(a) Raw Image



(b) Background Image



(c) Processed Image

Fig. 2 Image processing of PLIF images

nation process, which do not allow to directly relate the local fluorescence intensity to the local dye concentration.

The normalization of the raw images is based on the fact that the fluorescence intensity can be represented as the product of a quantity proportional to the local dye concentration and the local illumination such as

$$I_{\text{raw}}(x,y) \propto c_d(x,y) I_l(x,y),$$

where $I_{\text{raw}}(x,y)$ is the measured local fluorescence intensity, $c_d(x,y)$ is the local dye concentration and $I_l(x,y)$ is the local illumination intensity. In the case of the ensemble averaged background images obtained with a uniform dye concentration, the collected fluorescence intensity is essentially proportional to the local illumination map

$$\overline{I}_{\text{cal}}(x,y) \propto I_l(x,y),$$

where $\overline{I}_{\text{cal}}(x,y)$ is the intensity in the ensemble averaged background images. Hence, the individual images can be normalized as follows:

$$I_c(x,y) = C \frac{I_{\text{raw}}(x,y)}{\overline{I}_{\text{cal}}(x,y) + \varepsilon},$$

where $I_c(x,y)$ is the normalized image intensity, C is a normalization constant, and ε is a small positive constant added to avoid division by zero in areas where the illumination may be zero (ε was set to 0.5 out of 256 gray levels for the current processing). The normalization constant C was picked by examining and processing all images from all cycles and crank angles and selecting its value to maximize the dynamic range of the image. The optimized value for C was then kept fixed for all images to allow the comparison among the intensity maps across different cycles and crank angles. Figure 2(*b*) represents the ensemble averaged background intensity $\overline{I}_{\text{cal}}(x,y)$ at the same crank angle as Fig. 2(*a*). Figure 2(*c*) shows the normalized image after applying the normalization procedure. This figure clearly shows that the ''streaky'' non-uniformity present in the raw image is virtually eliminated and that correcting for the non-uniformity of the laser light sheet reveals a region of high concentration on the right side of the image that is hardly visible in the raw image.

Results

Three-Dimensional PTV. Figure 3 shows the phase-averaged, crank angle evolution of the flow field in the center tumble plane, at every 18 deg of crank angle increment from 36 deg After Top Dead Center (ATDC) until Bottom Dead Center (BDC). Each plot shows the in-plane velocity components as a vector map, superposed with iso-contour lines for the out-of-plane velocity component. A reference vector located in the bottom left corner represents the mean piston speed as an indicative velocity scale. The location of the valves and piston are schematically indicated on each plot to facilitate the easier interpretation of the flow field. The primary flow pattern observed in this center tumble plane is the development of a well identified tumbling vortex as early as at 54 deg ATDC. However, this vortex is initially small and located under the exhaust valve, and progressively grows in size and strength. Furthermore, the center of this tumbling vortex shifts from the upper left corner (under the exhaust valve) initially towards the piston and the opposite side of the cylinder (lower right corner) at BDC.

The three-dimensional PTV technique not only yields the velocity map in this plane but over the entire volume. The other planes are not shown here for brevity, but the resulting flow patterns in the rest of the cylinder will be summarized. In addition to the strongly organized tumbling vortex near the center of the cylinder, two well organized cross-flow patterns (cross-tumble eddies) can be observed on both sides of the center plane (symmetry plane). These three-dimensional flow patterns are fairly typical of other dual intake valve engines (3 or 4 valve engine configurations). Detailed description of such flow fields can be found in the reference [11]. In summary, it is important to bear in mind that (1) these intake flow fields are rapidly evolving and (2) they are highly three-dimensional.

Cycle-Resolved Instantaneous Concentrations Maps. The evolution of the dye concentration during the intake stroke in the center tumble plane is shown in Fig. 4 for one particular cycle. The crank angles of each image (processed as described above) correspond directly to the crank angles for which the velocity field was shown in Fig. 3, i.e., spanning the intake stroke from 36 deg ATDD until BDC, in 18 deg increments. The concentration evolution shown here is very typical of all the cycles that were analyzed. During the initial part of the intake stroke, the dye (fuel) is only starting to appear immediately below the intake valves. As a matter of fact, in many other cycles (not shown here), the dye often has not yet reached the center tumble plane at the earliest crank angle shown here (36 deg ATDC). A short instant later, a highly concentrated dye region appears below the exhaust valve. This corresponds to the dye entrained by the primary convection pattern. This is an eddy (or pair of eddies) located essentially inside the combustion chamber, and which is fed by the flow from the intake valves, following the contour of the combustion chamber and being turned below the exhaust valves by the opposite wall of the cylinder (and the piston face). This convected region of dye (fuel) is very intense, as it tags the initial fluid entering the cylinder which is dye (fuel)-rich. In the following pictures 4(*c*) and 4(*d*), the convective nature of the process by the initial tumbling vortex is very evident. At the later crank angles, Figs. 4(*e*), 4(*f*), and 4(*g*), the convective nature of the process is still identifiable, albeit complicated by the appearance of dye (fuel) in the center tumble plane from other regions of the cylinder (consistent with the three-dimensional cross flow eddies observed in the velocity field, which convect dye (fuel) towards the center plane as well). However, the region of highest concentration is still directly linked to the main advective tumble vortex. By the latter part of the intake stroke, Figs. 4(*h*) and 4(*i*), some of the dye (fuel) is advected along the cylinder wall towards the head, while some gets trapped in the recirculation of the main tumbling vortex. This description of the process (shown for one cycle in Fig. 4) only characterizes the main features of the concentration field, which are essentially present in most cycles. In addition, these pictures reveal the finer scale structure of the concentration field and the local non-homogeneities present during this overall advective mixing process. These local non-homogeneities are probably sharper than in the actual air/fuel vapor mixture owing to the much greater molecular diffusivity. However, many inhomogeneities are on a scale of the order of 3 to 15 mm, which is much larger than the typical diffusive length scale during this time interval (of the order of 1 mm). Hence, except for the sharpness of the concentration field gradients, the local non-homogeneities observed here are probably representative of the degree of mixedness encountered in actual engines during the intake stroke.

Crank-Angle Evolution of Phase-Averaged Concentrations and Cycle-to-Cycle r.m.s Fluctuations. To better quantify the repeatability of this advective mixing process, as well as its cycle-to-cycle variability, the results from 50 cycles of data were averaged at each crank-angle. The results of this phase averaging are shown in Fig. 5, again for the same crank angles as before. Of course, owing to the nature of the phase averaging process, all the smaller scale features of the concentration maps disappear. However, the overall (large-scale) inhomogeneity is very consistent from cycle-to-cycle and is clearly brought to light by the phase-averaging process. As discussed earlier, the overall advective nature of the dye (fuel) concentration is very clear in this Figure as the dye (fuel) rich region can clearly be traced as it is advected around the cylinder by the dominant tumble vortex.

On the average, the dye (fuel) concentration is far from being uniform at the end of the intake stroke. The dye (fuel) lean flow

entering into the cylinder during the latter part of the intake stroke is not mixed with the dye (fuel) which entered the cylinder during the initial part of the intake stroke and which is mostly convected along the wall under the intake valves. Under some carefully crafted circumstances, this advective feature of the fuel mixing can actually be beneficial under lean burning conditions, by timing the arrival of a fuel rich zone in the vicinity of the spark plug at the time of ignition (hence promoting a good ignition under near stoichiometric conditions, while keeping the rest of the cylinder relatively lean).

By nature of the phase-averaging process, these concentration maps can be somewhat misleading, as they do not address the repeatability of this process from cycle to cycle. To quantify this important aspect of the process, the point-by-point, cycle-to-cycle r.m.s. deviations from the mean were calculated. These r.m.s. maps are shown in Fig. 6 for the same corresponding crank angles.They indicate that certain regions, usually located near the edges of the high concentration regions (Ref. Fig. 5), are associated with substantial cycle-to-cycle fluctuations. Since the overall process is primarily advective, these cycle-to-cycle concentration variations are primarily associated with the ''jittering'' of well-formed velocity patterns (eddies). In other words, while the flow patterns themselves are repeatable, their location and strength as a function of time varies from cycle to cycle, hence contributing to substantial concentration variations near the ''average'' edge of the eddies. In addition, high r.m.s. regions are also noticeable along the piston crown and the bottom of the cylinder wall under the intake valve due to the occasional ''trapping'' of high dye

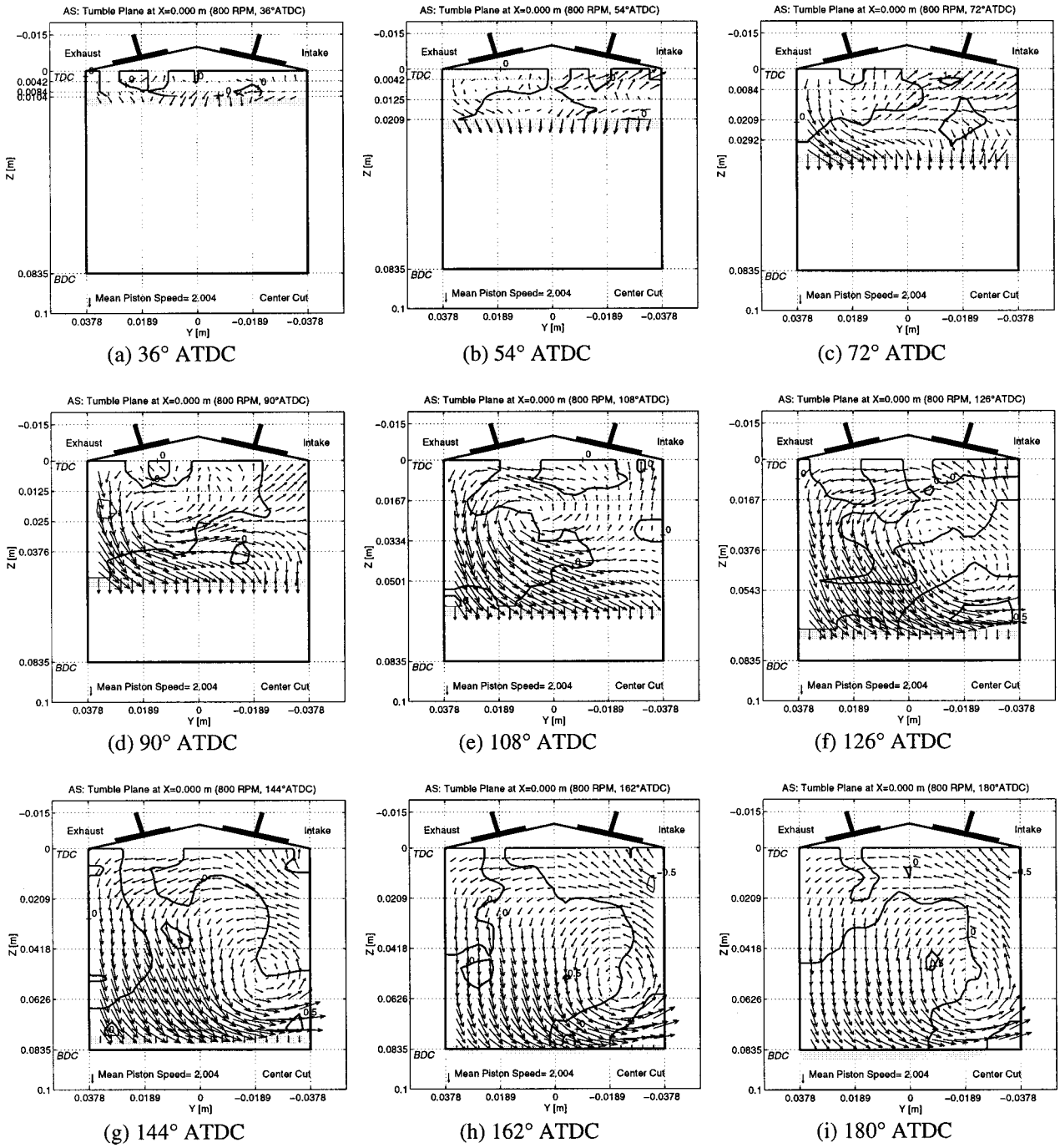


Fig. 3 Ensemble averaged velocity field in center tumble plane

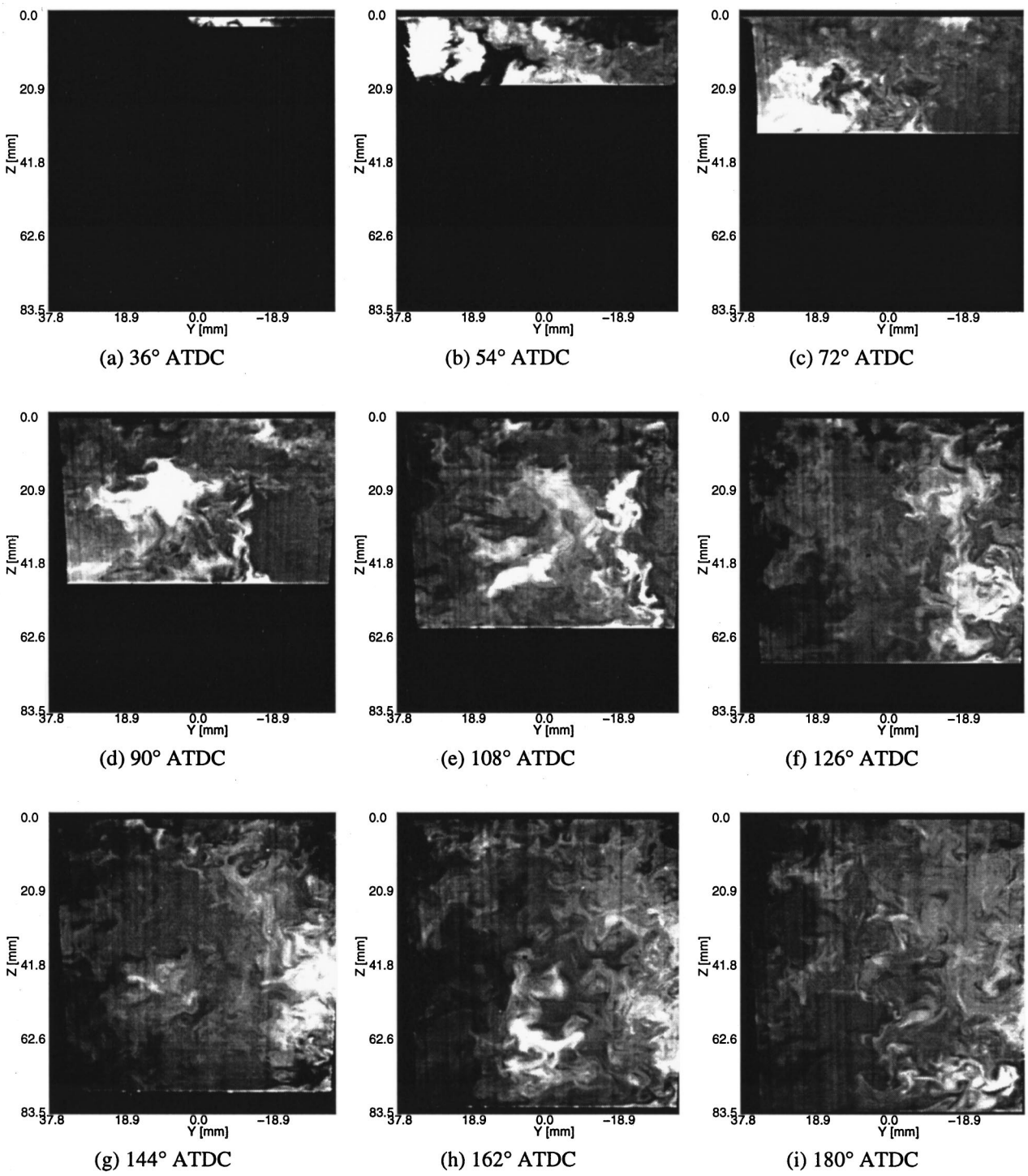


Fig. 4 Crank angle evolution of PLIF images during one cycle in center tumble plane

(fuel) concentration layers near the wall, away from the primary convective pattern (tumble and associated cross-tumble eddies). It should also be noted (not really visible in the gray scale images), that the high r.m.s. levels are almost of the same order as the high average concentration levels. In other words, the convective mixing process can give rise to significant cycle-to-cycle local variations, at least during the intake stroke.

Crank-Angle Evolution of Area-Averaged Concentration and Uniformity. Finally, overall area-averaged statistics of the concentration field were computed. Specifically, for each crank angle, the phase-averaged concentration maps were spatially averaged over the entire plane to provide an overall progression of the average dye (fuel) concentration in the entire cylinder in the center tumble plane. This is shown in Fig. 7(*a*), which depicts the crank-angle evolution of the average concentration level. It shows that the initial concentration rises rapidly as the dye (fuel)-rich fluid enters the cylinder during the early part of the intake stroke, peaking around 70 deg or 80 deg ATDC. Then, the fluid entering into the cylinder is essentially very lean and the overall concentration in the cylinder decreases and stabilizes as the piston ap-

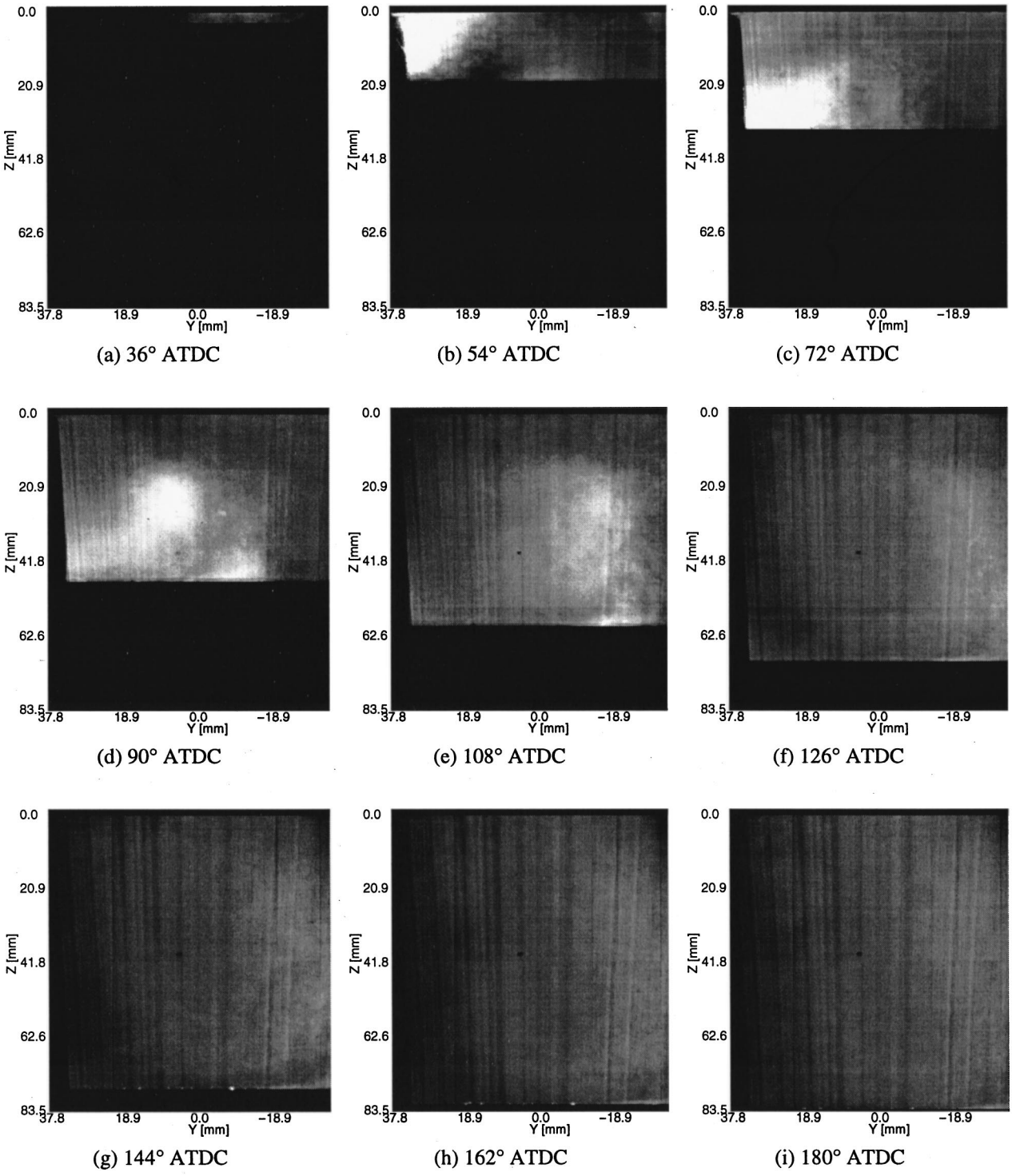


Fig. 5 Crank angle evolution of phase averaged PLIF images in center tumble plane

proaches BDC. This final concentration is essentially representative of the overall air/fuel ratio in an actual engine. In the arbitrary concentration units used here, this value is of the order of 23. This is to be contrasted to the local concentration values which range between essentially 0 or 5 (in the fuel lean regions) and 50 or 60 (in the fuel rich regions). To quantify this overall spatial non-uniformity, the spatial r.m.s. concentration fluctuations at each crank angle were also calculated and are shown in Fig. 7(*b*). Again, the spatial r.m.s. distribution peaks early in the intake stroke, as the dye (fuel) enters the cylinder as a well defined, concentrated region. As the convective mixing process takes place during the intake stroke, this spatial r.m.s. decreases continuously, even as BDC is approached, indicating that the convective mixing by the eddies does not depend only on the creation of new eddies during the induction process, but also the persistence of these formed flow patterns around BDC.

It should also be noted that the spatial non-uniformity levels (r.m.s. of the order of 20 units at BDC) is almost as large as the

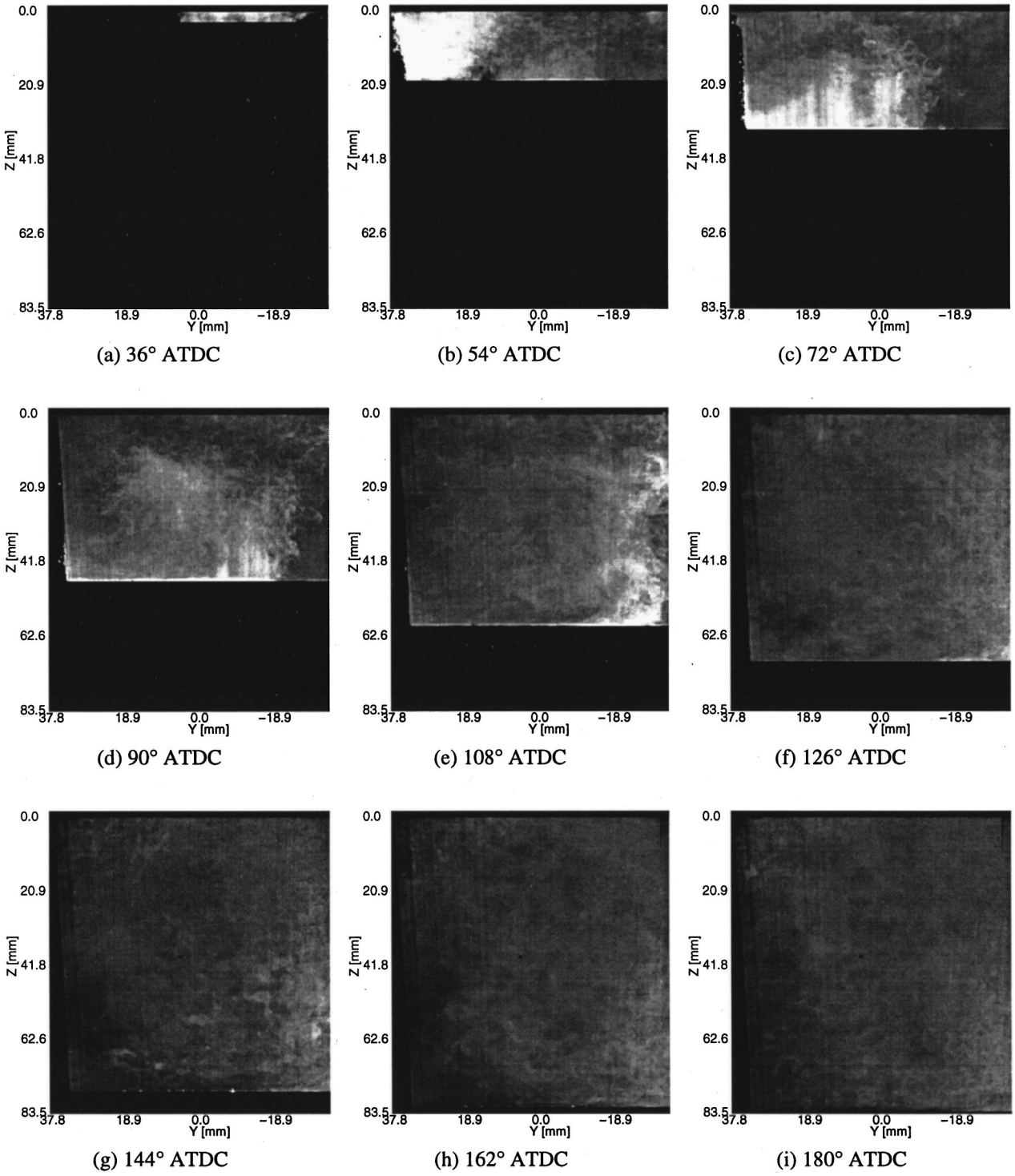
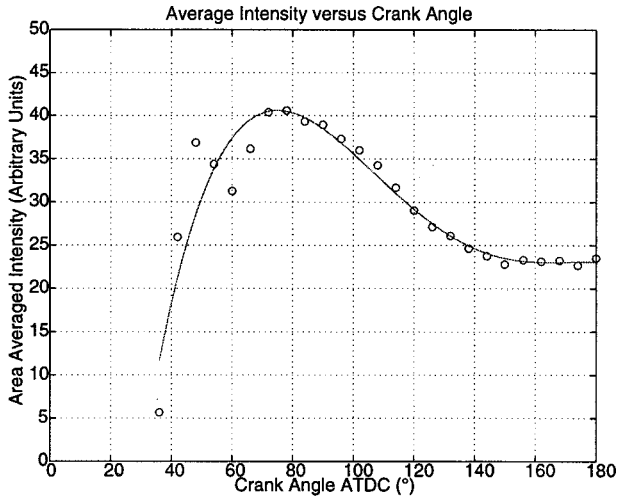


Fig. 6 Crank angle evolution of cycle-to-cycle RMS fluctuations of PLIF images in center tumble plane

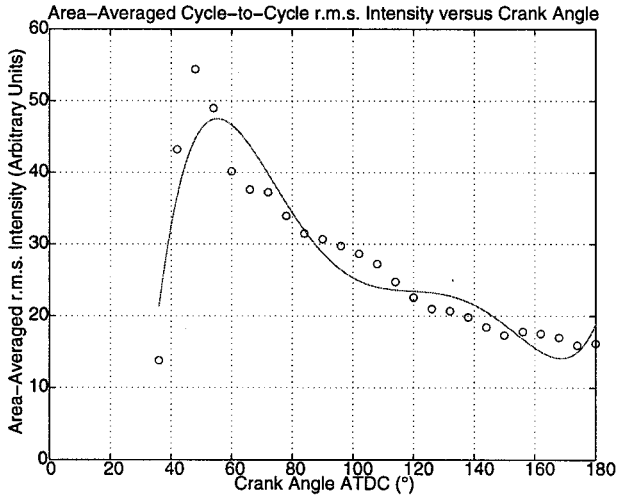
average concentration (of the order of 23 units), hence consistent with large scale, spatial non-uniformity existing at the end of the intake stroke.

Concluding Remarks

In summary, several points can be made. First, this paper demonstrates that a great deal of information about the advective mixing process dominating the intake stroke can be gathered by a water analog and aqueous dye-based LIF methodology especially for rapid evaluation of the relative merit analysis. The primary advantage of this approach is its relative ease of implementation, hence providing a valuable tool for engine development. While we perfectly recognize the limitations of this approach (cannot study the compression stroke, and cannot adequately capture the diffusive mixing process), the dominant flow physics governing the fuel distribution/mal-distribution in the cylinder during the intake stroke can be better understood and optimized. Furthermore, this approach is a natural counterpart to the detailed three-



(a) Average Intensity



(b) Cycle-to-Cycle r.m.s. Fluctuations

Fig. 7 Crank angle evolution of area-averaged intensity level and cycle-to-cycle r.m.s. fluctuations

dimensional flow measurements that are now routinely being carried out using a water analog rig and the three-dimensional Particle Tracking Velocimetry approach. Linking the two techniques in the same experiment, under identical operating (and simulation) conditions, provides great insight into the mechanistic nature of the initial mixing inside the cylinder.

Second, this study clearly underscores the importance of the in-cylinder flow patterns created during the intake stroke to control the subsequent combustion process. While a clear link between the strength of the flow pattern at the end of the intake stroke (BDC) and combustion parameters has been established for a large variety of engine configurations, the intake-generated flow patterns are typically viewed as a ''storage'' of kinetic energy to be used later. This ''stored'' kinetic energy is then used as a source of turbulent kinetic energy at the time of spark, hence promoting faster turbulent flame speed through an enhanced turbulent kinetic energy and the appropriate length scale. However, the concentration inhomogeneities (both at the large scale and at intermediate scales) observed at the end of the intake stroke also require the continued action of these large scale eddies. This continued advective mixing process through the compression stroke is required to further homogenize the fuel distribution on a scale over which molecular diffusive processes can act in the time scale governed by the duration of the compression stroke. In that sense, ''robust'', large scale flow patterns (i.e., eddies which do not dissipate too fast) are paramount during the early part of the compression stroke to continue the effective advective mixing.

References

[1] Seitzman, J. M., and Hanson, R. K., 1993, *Instrumentation for Flows with Combustion*, A. M. K. P. Taylor, ed., Academic Press, San Diego, pp. 405–466.

[2] Baritaud, T. A., and Heintze, T. A., 1992, ''Gasoline Distribution Measurements with PLIF in an I. C. Engine,'' SAE Paper 922355.

[3] Wolff, D., Schluter, H., Beushausen, V., and Andersen, P., 1993, ''Quantitative Determination of Fuel Air Mixture Distributions in an Internal Combustion Engine Using PLIF of Acetone,'' Ber. Bunsenges. Phys. Chem., **97**, pp. 1738–1741.

[4] Zhao, F.-Q., Taketomi, M., Nishido, K., and Hiroyasu, H., 1994, ''PLIF Measurement of the Cyclic Variation of the Mixture Concentration in an SI Engine,'' SAE Paper 940988.

[5] Neij, H., Johansson, B., and Alden, M., 1994, ''Development and Demonstration of 2D-LIF for Studies of Mixture Preparation in SI Engines,'' Combust. Flame, **99**, pp. 449–457.

[6] Khaligi, B., and Huebler, M. S., 1988, ''A Transient Water Analog of Dual Intake Valve Engine for Intake Flow Visualization and Full-Field Velocity Measurements,'' SAE Paper 880519.

[7] Kent, J. C., Mikulec, A., Rimai, L., Adamczyk, A. A., Mueller, S. R., Stein, R. A., and Warren, C. C., 1989, ''Observations on the Effects of Intake-Generated Swirl and Tumble on Combustion Duration,'' SAE Paper 892096.

[8] Jackson, N. S., Stokes, J., Heikle, M., and Downie, J., 1995, ''A Dynamic Water Flow Visualization Rig for Automotive Combustion Systems Development,'' SAE Paper 950728.

[9] Guezennec, Y. G., Brodkey, R. S., Trigui, N., and Kent, J. C., 1994, ''Algorithms for Fully-Automated Three-Dimensional Particle Tracking Velocimetry,'' Exp. Fluids, **17**, pp. 209–219.

[10] Kent, J. C., Trigui, N., Choi, W.-C., and Guezennec, Y. G., 1994, ''Characterization of Intake-Generated Flow Fields in an IC Engine Using 3-D Particle Tracking Velocimetry (3-D PTV),'' SAE Paper 940279.

[11] Choi, W.-C., Guezennec, Y. G., and Jung, I., 1996, ''Rapid Evaluation of Variable Valve Lift Strategies Using 3-D In-Cylinder Flow Measurements,'' SAE Paper 96095.

[12] Trigui, N., Choi, W.-C., and Guezennec, Y. G., 1996, ''Cycle-Resolved Turbulence Intensity Measurements in IC Engines,'' SAE Paper 962085.

[13] Denlinger, A. R., Guezennec, Y. G., and Choi, W.-C., 1998, ''Dynamic Evolution of the 3-D Flow Field During the Latter Part of the Intake Stroke in an IC Engine,'' SAE Paper 980485.

[14] Trigui, N., Kent, J. C., and Affes, H., 1994, ''Use of Experimentally-Measured in-Cylinder Flow Field at IVC as Initial Conditions to CFD Simulations of Compression Stroke Analysis,'' SAE Paper 940208.

[15] Kooshesfahani, M. M., and Dimotakis, P. E., 1986, ''Mixing and Chemical Reactions in a Turbulent Liquid Mixing Layer,'' J. Fluid Mech., **170**, pp. 83–112.

[16] Dahm, W. J. A., and Dimotakis, P. E., 1990, ''Mixing at Large Schmidt Number in the Self-Similar Far Field of Turbulent Jets,'' J. Fluid Mech., **217**, pp. 299–330.

[17] Merkel, G. J., Rys, P., Rys, F. S., and Dracos, T. A., 1995, ''Concentration and Velocity Field Measurements in Turbulent Flows by Laser-Induced Fluorescence Tomography,'' *Proc. 7th Int. Symp. Flow Visualization*, Seattle, WA.

[18] Buch, K. A., and Dahm, W. J. A., 1996, ''Experimental Study of the Fine-Scale Structure of Conserved Scalar Mixing in Turbulent Shear Flows: Part I—Sc>>1,'' J. Fluid Mech., **317**, pp. 21–71.

[19] Choi, W.-C., and Guezennec, Y. G., 1991, ''*In Situ* Calibration for Wide Angle 3-Dimensional Stereoscopic Image Analysis,'' Appl. Opt., **36**, pp. 7364–7373.